Single-Agent Stability in Additively Separable Hedonic Games *

Felix Brandt, Martin Bullinger*, Leo Tappe

School of Computation, Information and Technology
Technische Universität München, Boltzmannstr. 3, 85748 Munich, Germany

Abstract

The formation of stable coalitions is a central concern in multiagent systems. A considerable stream of research defines stability via the absence of beneficial deviations by single agents. While most of the literature focuses on deviations constrained by unanimous consent, we also study consent decided by majority vote, and introduce two new stability notions that can be seen as local variants of popularity. We investigate these notions in additively separable hedonic games by pinpointing boundaries to computational complexity depending on the type of consent and friend-oriented utility restrictions. Many of our positive results follow from a new combinatorial observation that we call the *Deviation Lemma* and that we leverage to prove the convergence of simple and natural single-agent dynamics under fairly general conditions. Our negative results cover in particular the complexity of contractual Nash stability.

Keywords: Computational Social Choice, Algorithmic Game Theory, Hedonic Games, Single-Agent Stability, Dynamics

1. Introduction

Coalition formation is a central concern in multi-agent systems and considers the question of grouping a set of agents, e.g., humans or machines, into coalitions such as teams, clubs, or societies. A prominent framework for studying coalition formation is that of hedonic games, where agents' utilities are solely based on the coalition they are part of, and which thus disregards inter-coalitional relationships (Drèze and Greenberg, 1980). Hedonic games have been successfully used to model various scenarios evolving from operations research or the mathematical social sciences, such as research team formation (Alcalde and Revilla,

*This paper unifies and expands results that appeared in the Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI) (Brandt et al., 2022) and the Proceedings of the 47th International Symposium on Mathematical Foundations of Computer Science (MFCS) (Bullinger, 2022).

*Corresponding author

2004), task allocation (Saad et al., 2011), or community detection (Aziz et al., 2019). Identifying desirable coalition structures is often based on the prospect of coalitions to stay together. To this end, various notions of stability have been introduced and studied. A coalition structure (henceforth partition) is stable when no individual or group of agents benefits by joining another coalition or by forming a new coalition.

In this paper, we focus on deviations by single agents. The simplest example is a Nash deviation where some agent unilaterally decides to leave her current coalition in order to join another coalition. While such a deviation clearly captures the incentive of single agents to perform deviations, it completely ignores the other agents' opinions about the deviation. To overcome this shortcoming, various restrictions of Nash deviations have been proposed. This has motivated stability notions, such as individual stability or contractual Nash stability, which consider the unanimous consent of some or all of the coalitions directly affected by the deviation. While unanimous consent is in fact used in the formation process of international bodies like the EU or the NATO, it might be impractical and even undesirable in small- or medium-scale coalition formation scenarios. As a compromise, we also study intermediate notions of stability based on majority votes among the involved coalitions. This setting has received little attention so far (Gairing and Savani, 2019), and we will also define new majority-based stability notions.

Since the number of coalitions an agent can be part of is not polynomially bounded, a lot of effort has been put into identifying reasonable and succinct classes of hedonic games (see, e.g., Aziz et al., 2019; Ballester, 2004; Bogomolnaia and Jackson, 2002; Elkind and Wooldridge, 2009). In many such classes, agents extract cardinal preferences from a weighted graph by some aggregation method. Perhaps the most natural and thoroughly studied way to aggregate preferences is by taking the sum of the weights of edges towards agents in one's own coalition. This leads to the concept of additively separable hedonic games (ASHGs) (Bogomolnaia and Jackson, 2002). ASHGs allow the modeling of settings where agents have friends and enemies, and their goal is to simultaneously maximize the number of friends and minimize the number of enemies, while one of these two goals can have higher priority than the other one (Dimitrov et al., 2006). Our work provides a computational analysis of single-agent stability in ASHGs with a focus on friend-oriented utility restrictions.

1.1. Contribution

A recent line of research on stability notions focuses on the dynamical aspects leading to the formation of stable outcomes (see, e.g., Bilò et al., 2018; Hoefer et al., 2018; Carosi et al., 2019; Brandt et al., 2021). This yields an important distributed perspective on the coalition formation process. The value of some positive computational results in the context of hedonic games is diminished by the fact that they implicitly assume that a central authority has the means to collect all individual preferences, compute a stable partition, and enforce this partition on the agents. In contrast, simple dynamics based on single-agent deviations provide a much more plausible explanation for the formation of stable partitions. A versatile tool to prove the convergence of dynamics are potential functions, which guide the dynamics towards stable states (see, e.g., Bogomolnaia and Jackson, 2002; Suksompong, 2015; Brandt et al., 2021; Bullinger and Suksompong, 2023).

We extend the applicability of this approach by considering *non-monotonic* potential functions, i.e., potential functions that might decrease in some rounds of the dynamic process. This is possible because the total number of rounds can be bounded by observing the potential function from a global perspective using a new general combinatorial insight that we call the Deviation Lemma. We demonstrate the power of this lemma by providing three applications, which in particular yield polynomial running time of dynamics in friend-oriented games for various stability concepts. The Deviation Lemma is not restricted to additively separable utilities or the specific type of single-agent deviations. For instance, the combinatorial relationship of the lemma also arises naturally in the analysis of deviation dynamics in classes of games beyond the scope of this paper, such as anonymous hedonic games (Bogomolnaia and Jackson, 2002). In fact, the lemma holds for every sequence of partitions such that each partition evolves from its predecessor by having one element move to another partition class. It establishes a relationship between the development of the sizes of coalitions involved in deviations to information solely based on the starting partition and the terminal partition of the sequence.

For the special case of symmetric utility functions, additively separable hedonic games are well understood: the standard notion of utilitarian social welfare represents an increasing potential function for the dynamics induced by Nash stability (Bogomolnaia and Jackson, 2002), but finding stable states (even under unanimous consent of the welcoming coalition) leads to PLS-complete problems (Gairing and Savani, 2019). As we will see, this implies worst-case exponential running time of the dynamics. By contrast, our results hold for restricted sets of *non-symmetric* utility functions and our computational boundaries lie between polynomial-time computability and NP-completeness. In fact, whenever we identify a potential function guaranteeing the existence of stable outcomes, we are also able to prove that, from any starting partition, the corresponding simple dynamics of single-agent deviations converges to a stable partition in a polynomial number of rounds.

In contrast to the positive results obtained by means of the Deviation Lemma, we also find strong computational boundaries. We obtain NP-hardness of the existence problem for Nash stability in severely restricted ASHGs as well as the existence problem of contractually Nash-stable coalition structures in general ASHGs. Despite knowing that additively separable hedonic games that do not admit a contractually Nash-stable coalition structure exist (Sung and Dimitrov, 2007), previous investigations of single-agent stability have left the complexity of the associated existence problem open (Sung and Dimitrov, 2010). Hence, we complete the picture of the complexity of unanimity-based single-agent stability concepts in ASHGs.

In addition, we also find computational boundaries for majority-based stability concepts. This complements the results obtained by the Deviation Lemma.

Together, we completely pinpoint the complexity of majority-based stability notions in appreciation-of-friends games, aversion-to-enemies games, and friends-and-enemies games. Notably, a major step towards these hardness results is the construction of No-instances, which then can be levered in hardness reductions.

Our results are in line with the repeatedly observed theme in hedonic games research that the existence of counterexamples is the key to computational intractability (see, e.g., Dimitrov et al., 2006; Sung and Dimitrov, 2010; Aziz et al., 2013; Brandt et al., 2021).[1] On the other hand, we demonstrate that the observed intractabilities lie at the computational boundary by carving out further weak restrictions that lead to the existence and efficient computability of stable states.

1.2. Related Work

The study of hedonic games was initiated by Drèze and Greenberg (1980) but was only popularized two decades later by Banerjee et al. (2001), Cechlárová and Romero-Medina (2001), and Bogomolnaia and Jackson (2002). Aziz and Savani (2016) provide an overview of many important concepts. Two important research questions concern the design of reasonable computationally manageable subclasses of hedonic games and the detailed investigation of their computational properties. The former has led to a broad landscape of game representations. Some of these representations are ordinal and fully expressive, i.e., they can, in principle, express every preference relation over coalitions (Ballester, 2004; Elkind and Wooldridge, 2009). Still, representing certain preference relations requires exponential space. These representations are contrasted by cardinal representations based on weighted graphs (Aziz et al., 2019; Bogomolnaia and Jackson, 2002; Olsen, 2012), which are not fully expressive but only require polynomial space (except when weights are artificially large). Apart from the already discussed additively separable hedonic games, important aggregation methods consider the average of weights leading to the classes fractional hedonic games (Aziz et al., 2019) and modified fractional hedonic games (Olsen, 2012).

The computational properties of hedonic games have been extensively studied and we focus on literature related to additively separable hedonic games. Various versions of stability have been investigated (Dimitrov et al., 2006; Sung and Dimitrov, 2010; Aziz and Brandl, 2012; Aziz et al., 2013; Gairing and Savani, 2019). In particular, Sung and Dimitrov (2010) perform a detailed computational study of single-agent stability. Gairing and Savani (2019) settle the complexity of single-agent stability for symmetric input graphs. Apart from stability, other desirable axioms concern efficiency and fairness. Aziz et al. (2013) cover a wide range of axioms, whereas Elkind et al. (2020) and Bullinger (2020) focus on Pareto optimality, and Brandt and Bullinger (2022) investigate popularity, an axiom combining ideas from stability and efficiency, which is also related to a majority-based stability notion that we will introduce.

[1]A notable exception is provided by Bullinger and Kober (2021) who identify a class of hedonic games where partitions in the core always exist, but are still hard to compute.

The dynamical aspects of the coalition formation process have been studied in a series of very recent papers (Bilò et al., 2018; Hoefer et al., 2018; Carosi et al., 2019; Fanelli et al., 2021; Brandt et al., 2021; Boehmer et al., 2023; Bullinger and Suksompong, 2023). Most related is the work by Bilò et al. (2018) who consider Nash stability in fractional hedonic games and by Brandt et al. (2021) who consider dynamics based on individual stability in several classes of hedonic games. Bullinger and Suksompong (2023) consider a generalization of additively separable hedonic games and a stability concept analogous to Nash stability. Hoefer et al. (2018); Carosi et al. (2019), and Fanelli et al. (2021) consider dynamics based on group deviations. Finally, very recently, Boehmer et al. (2023) propose a dynamical version of hedonic games where utilities are modified based on the history of the performed deviations. They study both single-agent and group stability. Similar dynamic processes have been studied in the domain of matchings (see, e.g., Roth and Vande Vate, 1990; Abeledo and Rothblum, 1995; Brandt and Wilczynski, 2019).

2. Preliminaries and Model

In this section we introduce hedonic games and stability concepts. We use the notation $[k] = \{1, \ldots, k\}$ for any positive integer k.

2.1. Hedonic Games

Throughout the paper, we consider settings with a set $N = [n]$ of n agents. The goal of coalition formation is to find a partition of the agents into different disjoint coalitions according to their preferences. A *partition* of N is a subset $\pi \subseteq 2^N$ such that $\bigcup_{C\in\pi} C = N$, and for every pair $C, D \in \pi$, it holds that $C = D$ or $C \cap D = \emptyset$. An element of a partition is called *coalition*, and given a partition π, we denote by $\pi(i)$ the coalition containing agent i. We refer to the partition π given by $\pi(i) = \{i\}$ for every agent $i \in N$ as the *singleton partition*, and to $\pi = \{N\}$ as the *grand coalition*.

Let $\mathcal{N}_i$ denote all possible coalitions containing agent i, i.e., $\mathcal{N}_i = \{C \subseteq N : i \in C\}$. A *hedonic game* is defined by a tuple $(N, \succsim)$, where N is an agent set and $\succsim = (\succsim_i)_{i\in N}$ is a tuple of weak orders $\succsim_i$ over $\mathcal{N}_i$ which represent the preferences of the respective agent i. Hence, agents express preferences only over the coalitions which they are part of without considering externalities. The strict part of an order $\succsim_i$ is denoted by $\succ_i$, i.e., $C \succ_i D$ if and only if $C \succsim_i D$ and not $D \succsim_i C$.

The generality of the definition of hedonic games gives rise to many interesting subclasses of games that have been proposed in the literature. Many of these classes rely on cardinal utility functions $v_i : N \to \mathbb{R}$ for every agent i. Following Bogomolnaia and Jackson (2002), an *additively separable hedonic game* (ASHG) (N, v) consists of an agent set N and a tuple $v = (v_i)_{i\in N}$ of utility functions $v_i : N \to \mathbb{R}$ such that $\pi(i) \succsim_i \pi'(i)$ if and only if $\sum_{j\in\pi(i)} v_i(j) \geq \sum_{j\in\pi'(i)} v_i(j)$. Clearly, ASHGs are a subclass of hedonic games, and we can assume without loss of generality that $v_i(i) = 0$ (or set the utility of an agent for herself to an arbitrary constant).

Every ASHG can be naturally represented by a complete directed graph $G = (N, E)$ with weight $v_i(j)$ on arc (i, j). An ASHG is called *symmetric* if $v_i(j) = v_j(i)$ for every pair of agents i and j, and it can then be represented by a complete undirected graph with weight $v_i(j)$ on edge $\{i, j\}$. There are various subclasses of ASHGs that allow a natural interpretation in terms of friends and enemies. An agent $j \in N$ is called a *friend* (or *enemy*) of agent $i \in N$ if $v_i(j) > 0$ (or $v_i(j) < 0$). An ASHG is called a *friends-and-enemies game* (FEG) if $v_i(j) \in \{-1, 1\}$ for every pair of agents $i, j \in N$ (Brandt et al., 2022). Further, following Dimitrov et al. (2006), an ASHG is called an *appreciation-of-friends game* (AFG) (or an *aversion-to-enemies game* (AEG)) if $v_i(j) \in \{-1, n\}$ (or $v_i(j) \in \{-n, 1\}$). In all of these games, agents pursue the objective to maximize their number of friends while minimizing their number of enemies. In the case of an FEG, these two goals have equal priority, while there is a strict priority for one of the goals in AFGs and AEGs. Based on the friendship of agents, we define the *friendship relation* (or *enemy relation*) as the subset $R \subseteq N \times N$ where $(i, j) \in R$ if and only if $v_i(j) > 0$ (or $v_i(j) < 0$).

2.2. Stability Based on Single-Agent Deviations

We want to study stability under single agents' incentives to deviate. A *single-agent deviation* performed by agent i transforms a partition π into a partition π' where $\pi(i) \neq \pi'(i)$ and, for all agents $j \neq i$, it holds that $\pi(j) \setminus \{i\} = \pi'(j) \setminus \{i\}$. We write $\pi \xrightarrow{i} \pi'$ to denote a single-agent deviation performed by agent i transforming partition π to partition π'.

We consider myopic agents whose rationale is to only engage in a deviation if it immediately makes them better off. A *Nash deviation* is a single-agent deviation performed by agent i making her better off, i.e., $\pi'(i) \succ_i \pi(i)$. Any partition in which no Nash deviation is possible is said to be *Nash-stable* (NS).

This concept of stability is very strong and comes with the drawback that only the preferences of the deviating agent are considered. Therefore, various refinements have been proposed which additionally require the consent of the abandoned and the welcoming coalition. For a compact representation, we introduce them via the notion of favor sets. Let $C \subseteq N$ be a coalition and $i \in N$ an agent. The *favor-in set* of C with respect to i is the set of agents in C (excluding i) that strictly favor having i inside C rather than outside, i.e., $F_{\text{in}}(C, i) = \{j \in C \setminus \{i\} \colon C \cup \{i\} \succ_j C \setminus \{i\}\}$. The *favor-out set* of C with respect to i is the set of agents in C (excluding i) that strictly favor having i outside C rather than inside, i.e., $F_{\text{out}}(C, i) = \{j \in C \setminus \{i\} \colon C \setminus \{i\} \succ_j C \cup \{i\}\}$.

An *individual deviation* (or *contractual deviation*) is a Nash deviation $\pi \xrightarrow{i} \pi'$ such that $F_{\text{out}}(\pi'(i), i) = \emptyset$ (or $F_{\text{in}}(\pi(i), i) = \emptyset$). Then, a partition is said to be *individually stable* (IS) or *contractually Nash-stable* (CNS) if it allows for no individual or contractual deviation, respectively. A related weakening of both stability concepts is contractual individual stability (CIS), based on deviations that are both individual and contractual deviations (Bogomolnaia and Jackson, 2002; Dimitrov and Sung, 2007).

While these stability concepts include agents affected by the deviation, they require unanimous consent, which might be unnecessarily strong in some settings. Based on this observation, we define several hybrid stability concepts where the possibility of a deviation by some agent is decided via majority votes of the involved agents. A Nash deviation $\pi \xrightarrow{i} \pi'$ is called a *majority-in deviation* (or *majority-out deviation*) if $|F_{\text{in}}(\pi'(i), i)| \geq |F_{\text{out}}(\pi'(i), i)|$ (or $|F_{\text{out}}(\pi(i), i)| \geq |F_{\text{in}}(\pi(i), i)|$). A single-agent deviation that is both a majority-in deviation and a majority-out deviation is called *separate-majorities deviation*. Similar to before, a partition is said to be *majority-in stable* (MIS), *majority-out stable* (MOS), or *separate-majorities stable* (SMS) if it allows for no majority-in, majority-out, or separate-majorities deviation, respectively. The concepts MIS and MOS are special cases of the voting-based stability notions by Gairing and Savani (2019) for a threshold of 1/2.

Finally, it is possible to relax SMS by performing one joint vote instead of two separate votes. A Nash deviation $\pi \xrightarrow{i} \pi'$ is called a *joint-majority deviation* if $|F_{\text{out}}(\pi(i), i)| + |F_{\text{in}}(\pi'(i), i)| \geq |F_{\text{in}}(\pi(i), i)| + |F_{\text{out}}(\pi'(i), i)|$. A partition is then called *joint-majority stable* (JMS) if it allows for no joint-majority deviations. JMS is particularly interesting as it is a natural local version of popularity, an axiom recently studied in the context of hedonic games (Gärdenfors, 1975; Cseh, 2017; Brandt and Bullinger, 2020).[2]

Also note that while CIS is a refinement of Pareto optimality, there is no logical relationship between other (majority-based) stability concepts and Pareto optimality. In particular, we denote the stability concepts based on single-agent deviations by $\mathcal{S}$, i.e., $\mathcal{S} = \{\text{NS}, \text{IS}, \text{CNS}, \text{CIS}, \text{MIS}, \text{MOS}, \text{SMS}, \text{JMS}\}$. A taxonomy of our and related solution concepts is provided in Figure 1. For a more concise notation, we refer to deviations with respect to stability concept $\alpha \in \mathcal{S}$ as α *deviations*, e.g., IS deviations for $\alpha = \text{IS}$.

All these stability concepts naturally induce dynamics where we choose some starting partition and obtain a successor partition by having some agent perform a deviation from the current partition. More precisely, given a stability concept $\alpha \in \mathcal{S}$, an execution of α *dynamics* is an infinite or finite sequence $(\pi_j)_{j\geq 0}$ of partitions and a corresponding sequence $(i_j)_{j\geq 1}$ of (deviating) agents such that $\pi_{j-1} \xrightarrow{i_j} \pi_j$ is an α deviation for every j. The partition π_0 is then called the *starting partition*. Given a hedonic game G, and a stability concept $\alpha \in \mathcal{S}$, we say that the dynamics *converges for starting partition* π_0 if every execution of the α dynamics on G with starting partition π_0 is finite. Additionally, the α dynamics *converges* on G if it converges for every starting partition.

Proving convergence of dynamics is a very natural way to prove the existence of stable states and underlines the robustness of the stability concept. It complements a static solution concept with a decentralized process to reach a

[2] Informally speaking, a partition is popular if there is no other partition preferred by a majority of the agents. JMS partitions can only be challenged by partitions evolving through Nash deviations.

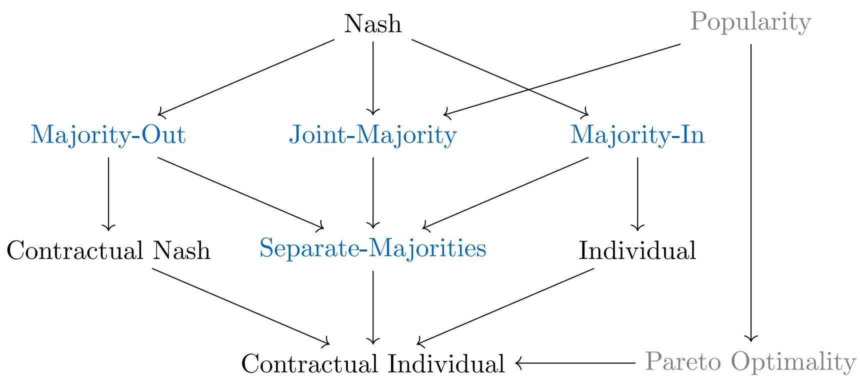


Figure 1: Logical relationships between stability notions and other solutions concepts. An arrow from concept α to concept β indicates that if a partition satisfies α, it also satisfies β. Majority-based stability notions are highlighted in blue, other single-agent based stability notions in black.

solution.

3. Computational Boundaries for Single-Agent Stability

In the next sections, we present our results. We start with computational boundaries for classical solution concepts.

First, we consider the notion of Nash stability. In the absence of negative utility values, the partition consisting solely of the grand coalition is Nash-stable. Conversely, in the absence of positive utility values, the singleton partition is Nash-stable. It is therefore necessary for an ASHG to have both positive and negative utility values in order to admit a non-trivial Nash-stable partition (see also Gairing and Savani, 2019).

Sung and Dimitrov (2010) showed that deciding whether an ASHG has an NS partition is NP-hard by a reduction from Exact Cover by 3-Sets. This reduction produces an ASHG with four distinct positive utility values and one negative utility value. We improve upon this result by showing that a reduction is possible with only one positive and one negative utility value. Moreover, it is possible for *any* choice of these two utility values, as long as the absolute value of the negative utility value is at least as large as the positive utility value. We state the theorem in a general way allowing the positive and negative utility value to be dependent on the number of agents of the particular instance. In this way, we simultaneously cover several important cases. For instance, the hardness holds for fixed constant positive and negative utility values as in FEGs, or for AFGs and AEGs. Note that for all of our stability notions, a stable partition is a polynomial-time verifiable certificate: one can simply check whether any agent can perform a deviation, and if no one can, the partition is stable. Therefore, we omit the proof of membership in NP in all of our reductions.

All of our hardness reductions are from the NP-complete problem Exact Cover by 3-Sets (E3C) (Karp, 1972). An instance of E3C consists of a tuple (R, S), where R is a ground set together with a set S of 3-element subsets of R. A Yes-instance is an instance such that there exists a subset $S' \subseteq S$ that partitions R. All omitted proofs can be found in the appendix.

Theorem 1. *Let $f^+ : \mathbb{N} \to \mathbb{Q}_{>0}$ and $f^- : \mathbb{N} \to \mathbb{Q}_{<0}$ be two polynomial-time computable functions satisfying $|f^-(m)| \geq f^+(m)$ for all $m \in \mathbb{N}$. Then, the problem of deciding whether an ASHG with utility values restricted to $\{f^-(n), f^+(n)\}$ has an NS partition is* NP*-complete.*

Theorem 1 requires the negative utility value to be at least as large in absolute value as the positive utility value. While we leave open the computational complexity for completely arbitrary pairs of negative and positive values, we can show that the problem is also hard when the positive utility value is significantly larger than the absolute value of the negative utility value. The reduction is a variant of the reduction in Theorem 1.

Theorem 2. *Deciding whether an AFG has an NS partition is* NP*-complete.*

Our next result settles the computational complexity of contractual Nash stability in ASHGs. Before giving the complete proof, we briefly describe the key ideas.

Given an instance (R, S) of Exact Cover by 3-Sets, the reduced instance consists of three types of gadgets. First, every element in R is represented by a subgame that does not contain a CNS partition. In principle, any such game can be used for a reduction, and we use a simple game identified by Sung and Dimitrov (2007). Moreover, we have further auxiliary gadgets that also consist of the same No-instance. The number of these auxiliary gadgets is equal to the number of sets in S that would remain after removing an exact cover of R, i.e., there are $|S| - |R|/3$ such gadgets. By design, the agents in the subgames corresponding to No-instances have to form coalitions with agents outside of their subgame in every CNS partition. The only agents that can achieve this are agents in gadgets corresponding to elements in S. A gadget corresponding to an element $s \in S$ can either prevent non-stability caused by exactly one auxiliary gadget, or by the three gadgets corresponding to the elements $r \in R$ with $r \in s$. Hence, the only possibility to deal with all No-instances simultaneously is if there exists an exact cover of R by sets in S. Then, the gadgets corresponding to elements in R can be dealt with by the cover and there are just enough elements in S to additionally deal with the other auxiliary gadgets.

Theorem 3. *Deciding whether an ASHG contains a CNS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let (R, S) be an instance of E3C and set $a = |S| - |R|/3$ (this is the number of additional sets in S if removing some exact cover). Without loss of generality, $a \geq 0$. We define an ASHG (N, v) as follows. Let $N = N_R \cup N_S \cup \bar{N}_S \cup N_A$ where

- $N_R = \cup_{r\in R} N_r$ with $N_r = \{r_i : i \in [4]\}$ for $r \in R$,
- $N_S = \cup_{s\in S} N_s$ with $N_s = \{s_r : r \in s\}$ for $s \in S$,
- $\bar{N}_S = \cup_{s\in S} \bar{N}_s$ with $\bar{N}_s = \{\bar{s}_r : r \in s\}$ for $s \in S$, and
- $N_A = \cup_{1\le j\le a} N^j$ with $N^j = \{x_i^j : i \in [4]\}$ for $1 \le j \le a$.

We define valuations v as follows:

- For each $r \in R$, $i \in [3]$: $v_{r_i}(r_4) = 1$.
- For each $r \in R$, $(i, j) \in (1, 2), (2, 3), (3, 1)$: $v_{r_i}(r_j) = 0$.
- For each $1 \le j \le a$, $i \in [3]$: $v_{x_i^j}(x_4^j) = 1$.
- For each $1 \le j \le a$, $(i, k) \in (1, 2), (2, 3), (3, 1)$: $v_{x_i^j}(x_k^j) = 0$.
- For each $s \in S$, $r \in s$: $v_{s_r}(r_4) = 1$.
- For each $s \in S$, $r \in s$, $1 \le j \le a$: $v_{s_r}(x_4^j) = v_{x_4^j}(s_r) = 0$.
- For each $s \in S$, $r, r' \in s$: $v_{s_r}(s_{r'}) = 0$.
- For each $s \in S$, $r, r' \in s$, $r \neq r'$, $z \in (N_S \cup N_A) \setminus N_s$: $v_{\bar{s}_r}(s_r) = 3$, $v_{\bar{s}_r}(s_{r'}) = -2$, and $v_{\bar{s}_r}(z) = 0$.
- All other valuations are -4.

An illustration of the game is given in Figure 2. The agents in N_R in the reduced instance form gadgets consisting of a subgame without CNS partition for every element in R. The agents in N_A constitute further such gadgets. The agents in N_S consist of triangles for every set in S and are the only agents who can bind agents in the gadgets in any CNS partition. Finally, agents in $\bar{N}_S$ avoid having agents in N_S in separate coalitions to bind agents in N_A.

We claim that (R, S) is a Yes-instance if and only if (N, v) contains a CNS partition.

$\Longrightarrow$. Suppose first that $S' \subseteq S$ partitions R. Consider any bijection $\phi\colon S \setminus S' \to [a]$. Define a partition π by taking the union of the following coalitions:

- For every $r \in R, i \in [3]$, form $\{r_i\}$.
- For $s \in S', r \in s$, form $\{s_r, r_4\}$.
- For $s \in S \setminus S'$, form $\{s_r : r \in s\} \cup \{x_4^{\phi(s)}\}$.
- For $s \in S, r \in s$, form $\{\bar{s}_r\}$.
- For $1 \le j \le a, i \in [3]$, form $\{x_i^j\}$.

We claim that π is CNS. We will show that no agent can perform a deviation.

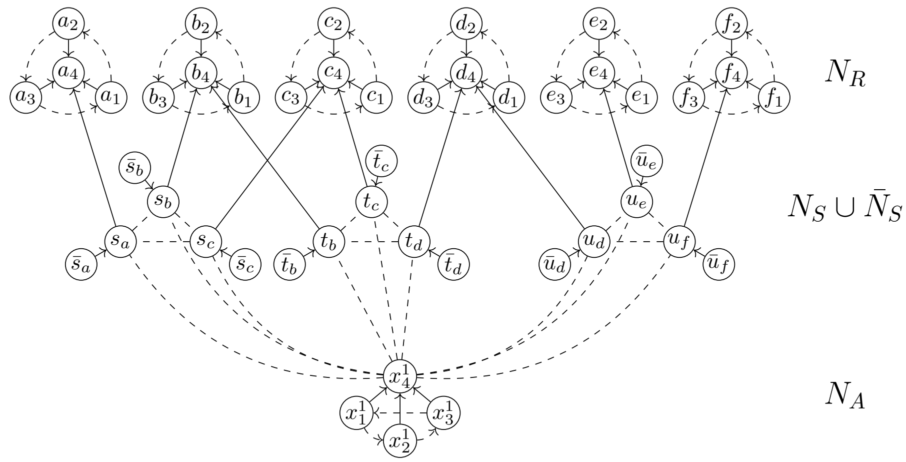


Figure 2: Schematic of the reduction from the proof of Theorem 3. We depict the reduced instance for the instance (R, S) of E3C where $R = \{a, b, c, d, e, f\}$, and $S = \{s, t, u\}$, with $s = \{a, b, c\}$, $t = \{b, c, d\}$, and $u = \{d, e, f\}$. Fully drawn edges mean a positive utility, which is usually 1 except between agents of the types $\bar{s}_r$ and s_r, where $v_{\bar{s}_r}(s_r) = 3$. Dashed edges represent a utility of 0. For agents in $\bar{N}_S$, only the single positive utility is displayed. Other omitted edges represent a negative utility of -4.

- For $r \in R$, $i \in [3]$, it holds that $v_{r_i}(\pi) = 0$ and joining any other coalition results in a negative utility. In particular, $v_{r_i}(\pi(r_4) \cup \{r_i\}) = -3$.
- For $r \in R$, r_4 is not allowed to leave her coalition.
- For $s \in S'$, $r \in s$, it holds that $v_{s_r}(\pi) = 1$ and joining any other coalition results in a negative utility. The agent s_r is in a most preferred coalition.
- For $s \in S \setminus S'$, $r \in s$, it holds that $v_{s_r}(\pi) = 0$ and joining any other coalition results in a negative utility. In particular, $v_{s_r}(\pi(r_4) \cup \{s_r\}) = -3$.
- For $s \in S'$, $r \in s$, the agent $\bar{s}_r$ obtains a non-positive utility by joining any other coalition. In particular, $v_{\bar{s}_r}(\pi(s_r) \cup \{\bar{s}_r\}) = -1$.
- For $s \in S \setminus S'$, $r \in s$, the agent $\bar{s}_r$ obtains a non-positive utility by joining any other coalition. In particular, $v_{\bar{s}_r}(\pi(s_r) \cup \{\bar{s}_r\}) = -1$.
- For $1 \leq j \leq a$, $i \in [3]$, it holds that $v_{x_i^j}(\pi) = 0$ and joining any other coalition results in a negative utility. In particular, $v_{x_i^j}(\pi(x_4^j) \cup \{x_i^j\}) = -11$.
- For $1 \leq j \leq a$, x_4^j is in a best possible coalition (achieving utility 0).

$\Longleftarrow$. Conversely, assume that (N, v) contains a CNS partition π. Define $S' = \{s \in S \colon \pi(s_r) \cap N_R \neq \emptyset \text{ for some } r \in s\}$. We will show first that S' covers all elements in R and then show that $|S'| = |R|/3$.

Let $r \in R$. Then, for all $i \in [3]$, $\pi(r_i) \subseteq N_r$. This follows because there is no agent who favors r_i in her coalition. Therefore, she would leave any coalition with an agent outside N_r to receive non-negative utility in a singleton coalition. Further, if there is no $s \in S$ with $r \in s$ such that $r_4 \in \pi(r_s)$, then $\pi(r_4) \subseteq N_r$. Indeed, if r_4 forms any coalition except a singleton coalition, she will receive negative utility, and then there must exist an agent who favors her in the coalition. Consequently, if $r_4 \notin \pi(r_s)$ for all $s \in S$ with $r \in s$, then r_4 is in a singleton coalition, or there exists $i \in [3]$ with $r_4 \in \pi(r_i)$, for which we already know that $\pi(r_i) \subseteq N_r$.

Assume now that $\pi(r_4) \subseteq N_r$. For $i, i' \in [3]$, $r_i \notin \pi(r_{i'})$ because then one of them would receive a negative utility and could perform a CNS deviation to form a singleton coalition. If $\{r_4\} \in \pi$, then r_1 would deviate to join her. Hence, there exists exactly one $i \in [3]$ with $\{r_i, r_4\} \in \pi$. Suppose without loss of generality that $\{r_1, r_4\} \in \pi$. But then, r_3 would perform a CNS deviation to join them, a contradiction. We can conclude that there exists $s \in S$ with $r \in s$ such that $r_4 \in \pi(r_s)$. Hence, $s \in S'$ and we have shown that S' covers R.

To bound the cardinality of S', we will show that, for every $1 \leq j \leq a$, there exists $s \in S \setminus S'$ with $N_s \subseteq \pi(x_4^j)$. Let therefore $1 \leq j \leq a$ and let $C = \pi(x_4^j)$. Similar to the considerations about agents in N_r, we know that $\pi(x_i^j) \subseteq X^j$ for $i \in [3]$, and that it cannot happen that $C \subseteq X^j$, and therefore $C \cap X^j = \{x_4^j\}$. In particular, there must be an agent $y \in N \setminus X^j$ with $y \in C$. Since no agent in C favors x_4^j to be in her coalition, we know that $v_{x_4^j}(\pi) \geq 0$ and therefore $C \subseteq \{x_4^j\} \cup N_S$. Let $s \in S$ and $r \in s$ with $s_r \in C$. As we already know that $\bar{s}_r \notin C$, it must hold that $N_s \subseteq C$ to prevent her from joining. It follows that $s \notin S'$. Since $\pi(x_4^j) \cap \pi(x_4^{j'}) = \emptyset$ for $1 \leq j' \leq a$ with $j' \neq j$, we find an injective mapping $\phi \colon [a] \to S \setminus S'$ such that, for every $1 \leq j \leq a$, $N_{\phi(j)} \subseteq \pi(x_4^j)$. Consequently, $|S'| \leq |S| - |\phi([a])| \leq |S| - a = |R|/3$. Hence, S' covers all elements from R with (at most) $|R|/3$ sets and therefore is an exact cover. $\square$

The reduction in the previous proof only uses a very limited number of different weights, namely the weights in the set $\{1, 0, -2, -4\}$, where the weight -4 may be replaced by an arbitrary smaller weight. By contrast, we will see that CNS partitions always exist if the utility functions of an ASHG assume at most one nonpositive value, and can be computed efficiently in this case (cf. Theorem 5). This encompasses for instance FEGs, AFGs, and AEGs. Hence, the hardness result is close to the boundary of computational feasibility.

4. Deviation Lemma and Applications

By contrast, restricting the utility values to one positive and one negative value leads to positive results for other notions of stability. These results can be shown in a unified manner using a potential function argument that crucially hinges on the following general observation.

Lemma 1 (Deviation Lemma). *Let $\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k$ be a sequence of k*

single-agent deviations. Then, the following identity holds:

$$\sum_{j\in[k]}|\pi_j(i_j)|-|\pi_{j-1}(i_j)| = \frac{1}{2}\sum_{i\in N}|\pi_k(i)|-|\pi_0(i)|. \tag{1}$$

Proof. Let $\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k$ be a sequence of k single-agent deviations and fix some $j \in [k]$. Then, the following facts hold:

$$\begin{aligned}
|\pi_j(i_j)| &= \left(\sum_{i\in\pi_j(i_j)\setminus\{i_j\}}|\pi_j(i)|-|\pi_{j-1}(i)|\right)+1,\\
|\pi_{j-1}(i_j)| &= \left(\sum_{i\in\pi_{j-1}(i_j)\setminus\{i_j\}}|\pi_{j-1}(i)|-|\pi_j(i)|\right)+1,\\
\pi_j(i) &= \pi_{j-1}(i) \quad \forall i\in N\setminus\left(\pi_j(i_j)\cup\pi_{j-1}(i_j)\right).
\end{aligned}$$

Combining these facts allows us to express the difference of the deviator's coalition sizes as follows:

$$\begin{aligned}
|\pi_j(i_j)|-|\pi_{j-1}(i_j)| = &\left(\sum_{i\in\pi_j(i_j)\setminus\{i_j\}}|\pi_j(i)|-|\pi_{j-1}(i)|\right)\\
&-\left(\sum_{i\in\pi_{j-1}(i_j)\setminus\{i_j\}}|\pi_{j-1}(i)|-|\pi_j(i)|\right)\\
&+\sum_{i\in N\setminus(\pi_j(i_j)\cup\pi_{j-1}(i_j))}|\pi_j(i)|-|\pi_{j-1}(i)|\\
&=\sum_{i\in N\setminus\{i_j\}}|\pi_j(i)|-|\pi_{j-1}(i)|.
\end{aligned}$$

Adding $|\pi_j(i_j)|-|\pi_{j-1}(i_j)|$ to both sides yields

$$2\left(|\pi_j(i_j)|-|\pi_{j-1}(i_j)|\right) = \sum_{i\in N}|\pi_j(i)|-|\pi_{j-1}(i)|.$$

Summing these terms for all $j \in [k]$, interchanging summation order, and telescoping gives

$$\begin{aligned}
\sum_{j\in[k]}2\left(|\pi_j(i_j)|-|\pi_{j-1}(i_j)|\right) &= \sum_{j\in[k]}\sum_{i\in N}|\pi_j(i)|-|\pi_{j-1}(i)|\\
2\sum_{j\in[k]}|\pi_j(i_j)|-|\pi_{j-1}(i_j)| &= \sum_{i\in N}\sum_{j\in[k]}|\pi_j(i)|-|\pi_{j-1}(i)|\\
2\sum_{j\in[k]}|\pi_j(i_j)|-|\pi_{j-1}(i_j)| &= \sum_{i\in N}|\pi_k(i)|-|\pi_0(i)|.
\end{aligned}$$

Dividing both sides by 2 completes the proof. □

The Deviation Lemma is especially useful as the right-hand side of Equation (1) does not depend on k, and we can therefore also find bounds for its left-hand side solely depending on the number of players n.

Lemma 2. *Consider a sequence of k successive single-agent deviations*

$$\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k.$$

Then, the following bounds hold:

$$-\frac{n(n-1)}{2} \leq \sum_{j\in[k]} |\pi_j(i_j)| - |\pi_{j-1}(i_j)| \leq \frac{n(n-1)}{2}.$$

Proof. Observe that for all $i \in N$ and all partitions π, we have

$$1 \leq |\pi(i)| \leq n.$$

Thus, we can find the bounds

$$-n(n-1) \leq \sum_{i\in N} |\pi_k(i)| - |\pi_0(i)| \leq n(n-1).$$

Applying Lemma 1 yields the desired result. □

We demonstrate the power of the Deviation Lemma by proving convergence of the dynamics for a variety of deviation types and classes of ASHGs.

Theorem 4. *The dynamics of IS deviations always converges in ASHGs with at most one nonnegative utility value.*

Proof. Let (N, v) be an ASHG such that the v_i take on at most one nonnegative value. If there are no nonnegative valuations, all IS deviations are singleton formations, so after at most n deviations, we reach a stable partition. Now, suppose that there is exactly one nonnegative utility value $x \geq 0$. If there are no negative valuations, then in case $x = 0$ we terminate immediately, and in case $x > 0$ the grand coalition will form after at most n^2 deviations. The latter holds because every deviation increases the number of pairs of agents which are part of the same coalition. Thus, we will now assume that in addition to the single nonnegative utility value x, there is at least one negative utility value, and we denote the largest absolute value of a negative utility value by y. Further, define $\Delta = \min\{v_i(C) - v_i(C') \colon i \in N,\, C, C' \in \mathcal{N}_i,\, v_i(C) > v_i(C')\}$. Intuitively, $\Delta > 0$ is the minimum improvement any agent is guaranteed to have when making a NS deviation. Further, consider the potential function Φ defined by the social welfare of a partition as $\Phi(\pi) = \sum_{i\in N} v_i(\pi)$.

Let us investigate how this potential changes for a single IS deviation $\pi \xrightarrow{i} \pi'$.

$$
\begin{aligned}
\Phi(\pi') - \Phi(\pi) &= \underbrace{v_i(\pi') - v_i(\pi)}_{\text{deviator}} \\
&+ \underbrace{\sum_{j \in \pi'(i) \setminus \{i\}} v_j(\pi') - v_j(\pi)}_{\text{welcoming coalition}} + \underbrace{\sum_{j \in \pi(i) \setminus \{i\}} v_j(\pi') - v_j(\pi)}_{\text{abandoned coalition}} \\
&= v_i(\pi') - v_i(\pi) + \sum_{j \in \pi'(i) \setminus \{i\}} v_j(i) - \sum_{j \in \pi(i) \setminus \{i\}} v_j(i) \\
&= v_i(\pi') - v_i(\pi) + x\left(|\pi'(i)| - 1\right) - \sum_{j \in \pi(i) \setminus \{i\}} v_j(i) \\
&\geq \Delta + x\left(|\pi'(i)| - 1\right) - x\left(|\pi(i)| - 1\right) \\
&= \Delta + x\left(|\pi'(i)| - |\pi(i)|\right).
\end{aligned}
$$

The third equality comes from the fact that i performs an IS deviation, so all agents $j \in \pi'(i) \setminus \{i\}$ must accept i, which means they must have $v_j(i) = x$. Now, let π_0 be any initial partition and consider any sequence of k successive IS deviations

$$
\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k.
$$

Telescoping and termwise application of the above inequality yields $\Phi(\pi_k) - \Phi(\pi_0) = \sum_{j \in [k]} \Phi(\pi_j) - \Phi(\pi_{j-1}) \geq \sum_{j \in [k]} \Delta + x\left(|\pi_j(i_j)| - |\pi_{j-1}(i_j)|\right) = k\Delta + x \sum_{j \in [k]} |\pi_j(i_j)| - |\pi_{j-1}(i_j)|$. We recognize the sum from the Deviation Lemma, which can be bounded from below using Lemma 2:

$$
\Phi(\pi_k) - \Phi(\pi_0) \geq k\Delta - x\frac{n(n-1)}{2}. \tag{2}
$$

As the right hand side is unbounded in k, the sequence must be finite. To be precise, we can bound the potentials of the initial and final partitions by

$$
\Phi(\pi_0) \geq -n(n-1)y, \quad \Phi(\pi_k) \leq n(n-1)x.
$$

Substituting in these bounds and rearranging for k gives

$$
k \leq \frac{(2y + 3x)n(n-1)}{2\Delta}. \tag{3}
$$

□

There are a few important insights gained by the previous proof. First, the bound obtained via the Deviation Lemma does not mean that the potential function Φ is increasing in every round. In fact, since utilities are not necessarily symmetric, the deviating agent might move from a rather large coalition to a smaller coalition only improving her utility by Δ whereas the utility of all agents in the abandoned coalition are decreased by x. In fact, the Deviation Lemma does not give us control of the potential function in a single round. Also, it does not control the utility changes caused by the deviator. We apply it to control the utility changes of agents involved in deviations except for the deviator to obtain Equation (2). Hence, we can bound their utility changes by a global constant solely depending on input data. The utility changes caused by the deviator will then eventually lead to the potential reaching a local maximum.

Second, we can easily obtain polynomial bounds on the running time of the dynamics. If x and y are polynomially bounded in n and all valuations are integer, polynomial running time is directly obtained from Equation (3). In particular, this is the case for FEGs, AFGs, and AEGs, so individually stable partitions can be found in polynomial time for these games. After showing two more applications of the Deviation Lemma for other types of deviations, we will capture this observation in Corollary 1.

Third, the previous theorem is tight in the sense that the dynamics can cycle if we have two nonnegative utility values. Indeed, in an instance with agent set $N = [3]$ and utility values $v_i(j) = 1, v_j(i) = 0$ for $(i, j) \in \{(1, 2), (2, 3), (3, 1)\}$, the dynamics can infinitely cycle among the partitions $\{\{1, 2\}, \{3\}\}$, $\{\{1\}, \{2, 3\}\}$, and $\{\{1, 3\}, \{2\}\}$. However, the partition consisting of the grand coalition is individually stable and can be reached through the dynamics.

Our next application of the Deviation Lemma considers contractual Nash stability, where we obtain a similar result if we allow at most one nonpositive value. The proof is completely analogous and is therefore omitted. Note that this result also breaks down if we simultaneously allow the utility values -1 and 0 by constructing a similar cycle as in the previous example.

Theorem 5. *The dynamics of CNS deviations always converges in ASHGs with at most one nonpositive utility value.*

Theorems 4 and 5 use the Deviation Lemma to derive positive results for the single-sided unanimity-based stability notions IS and CNS. In a third application of the deviation lemma, we show that this technique is also applicable to majority-based stability notions, at least when we involve both the welcoming and the abandoned coalition in the vote. The key idea is a suitable arrangement of the terms occurring in the difference of the potential with respect to the agents affected by a deviation.

Theorem 6. *The dynamics of JMS deviations always converges in ASHGs with at most two distinct utility values.*

Proof. Let (N, v) be an ASHG such that the v_i take on at most two distinct values, and consider once again the potential

$$\Phi(\pi) = \sum_{i \in N} v_i(\pi).$$

If the v_i take on only one value or both values are nonnegative (resp., nonpositive), convergence is clear, as Φ increases with every JMS deviation. So suppose that the v_i are restricted to $\{-y, x\}$ with $y > 0$ and $x > 0$. As in the proof of Theorem 4, set $\Delta = \min\{v_i(C) - v_i(C') \colon i \in N,\, C, C' \in \mathcal{N}_i,\, v_i(C) > v_i(C')\}$.

Let us now investigate a single JMS deviation $\pi \xrightarrow{i} \pi'$. To reduce notational clutter, set $F_{\text{in}} = F_{\text{in}}(\pi(i), i), F_{\text{out}} = F_{\text{out}}(\pi(i), i), F'_{\text{in}} = F_{\text{in}}(\pi'(i), i)$, and $F'_{\text{out}} = F_{\text{out}}(\pi'(i), i)$. Note that, by definition of a JMS deviation, we have $|F'_{\text{in}}| + |F_{\text{out}}| \geq |F'_{\text{out}}| + |F_{\text{in}}|$, from which we can conclude

$$|F'_{\text{in}}| - |F_{\text{in}}| \geq \frac{|F'_{\text{in}}| - |F_{\text{in}}| + |F'_{\text{out}}| - |F_{\text{out}}|}{2} \geq |F'_{\text{out}}| - |F_{\text{out}}|.$$

Further, note that due to restriction of the utility values to $\{-y, x\}$, we have

$$\forall j \in F_{\text{in}} \cup F'_{\text{in}} : v_j(i) = x, \forall j \in F_{\text{out}} \cup F'_{\text{out}} : v_j(i) = -y$$

and

$$|F_{\text{in}}| + |F_{\text{out}}| = |\pi(i)| - 1, \quad |F'_{\text{in}}| + |F'_{\text{out}}| = |\pi'(i)| - 1.$$

Combining with our inequality from above, we obtain

$$|F'_{\text{in}}| - |F_{\text{in}}| \geq \frac{|\pi'(i)| - |\pi(i)|}{2} \geq |F'_{\text{out}}| - |F_{\text{out}}|.$$

The change in Φ through the JMS deviation can then be bounded as

$$\begin{aligned}
&\Phi(\pi') - \Phi(\pi) = \underbrace{v_i(\pi') - v_i(\pi)}_{\text{deviator}} \\
&+ \underbrace{\sum_{j \in \pi'(i) \setminus \{i\}} v_j(\pi') - v_j(\pi)}_{\text{welcoming coalition}} + \underbrace{\sum_{j \in \pi(i) \setminus \{i\}} v_j(\pi') - v_j(\pi)}_{\text{abandoned coalition}} \\
&= v_i(\pi') - v_i(\pi) + \sum_{j \in \pi'(i) \setminus \{i\}} v_j(i) - \sum_{j \in \pi(i) \setminus \{i\}} v_j(i) \\
&= v_i(\pi') - v_i(\pi) + x|F'_{\text{in}}| - y|F'_{\text{out}}| - x|F_{\text{in}}| + y|F_{\text{out}}| \\
&= v_i(\pi') - v_i(\pi) + x\left(|F'_{\text{in}}| - |F_{\text{in}}|\right) - y\left(|F'_{\text{out}}| - |F_{\text{out}}|\right) \\
&\geq \Delta + x\frac{|\pi'(i)| - |\pi(i)|}{2} - y\frac{|\pi'(i)| - |\pi(i)|}{2}.
\end{aligned}$$

Now, let π_0 be any initial partition and consider any sequence of k successive JMS deviations

$$\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \ldots \xrightarrow{i_k} \pi_k.$$

Telescoping and termwise application of the above inequality gives

$$\begin{aligned}
&\Phi(\pi_k) - \Phi(\pi_0) = \sum_{j \in [k]} \Phi(\pi_j) - \Phi(\pi_{j-1}) \\
&\geq \sum_{j \in [k]} \Delta + x\frac{|\pi_j(i_j)| - |\pi_{j-1}(i_j)|}{2} - y\frac{|\pi_j(i_j)| - |\pi_{j-1}(i_j)|}{2} \\
&= k\Delta + \frac{x - y}{2} \sum_{j \in [k]} |\pi_j(i_j)| - |\pi_{j-1}(i_j)|.
\end{aligned}$$

The sum from Lemma 1 appears for prefactors of different sign, and can be bounded using Lemma 2:

$$\begin{aligned}\Phi(\pi_k) - \Phi(\pi_0) &\geq k\Delta - \frac{x+y}{2}\frac{n(n-1)}{2} \\ &= k\Delta - \frac{(x+y)n(n-1)}{4}.\end{aligned}$$

As the right hand side is unbounded in k, the sequence must be finite. To be precise, we can bound the potentials of the initial and final partitions by

$$\Phi(\pi_0) \geq -n(n-1)y, \quad \Phi(\pi_k) \leq n(n-1)x.$$

Substituting in these bounds and rearranging for k gives

$$k \leq \frac{(5x+5y)n(n-1)}{4\Delta}.$$

□

Note that since every JMS deviation is also an SMS deviation, the previous result holds for SMS as well. As in the discussion after Theorem 4, we obtain a polynomial running time of the dynamics for appropriate restrictions of the cases. We collect important consequences in the following corollary. In particular, we extend results by Dimitrov et al. (2006) and Aziz and Brandl (2012) who proved the existence of IS partitions for AFGs and AEGs, respectively.[3]

Corollary 1. *The dynamics of IS, CNS, and JMS deviations always converges in polynomial time in AFGs, AEGs, and FEGs.*

We would like to stress that convergence of the dynamics does not guarantee a polynomial running time in general. An example is the case of symmetric utility values in ASHGs. For NS this can be directly inferred from the PLS reduction by Gairing and Savani (2019), which satisfies *tightness*, a property of reductions defined by Schäffer and Yannakakis (1991).

Proposition 1. *The dynamics of NS deviations in symmetric ASHGs may require exponentially many rounds before converging to an NS partition.*

Proof. It is easy to verify that the PLS reduction from PartyAffiliation under the Flip neighborhood by Gairing and Savani (2019, Observation 2) is tight. Schäffer and Yannakakis (1991, Lemma 3.3) showed that tight reductions preserve the existence of exponentially long running times of the standard local search algorithm, i.e., the NS dynamics in our case. Note that the standard local search algorithm of the source problem can have an exponential running time, because PartyAffiliation is a generalization of MaxCut whose standard

[3]These contributions actually show existence of partitions satisfying properties stronger than IS.

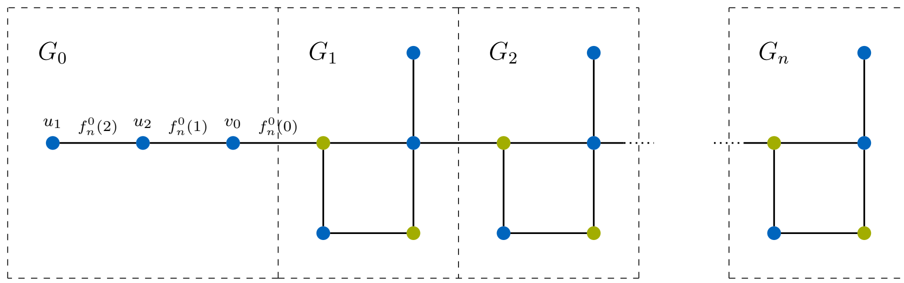


Figure 3: Exponential length IS dynamics inspired by Monien and Tscheuschner (2010). The starting partition into two coalitions is indicated by the coloring of the vertices.

local search algorithm can run in exponential time with respect to the flip neighborhood (Schäffer and Yannakakis, 1991, Theorem 5.15).[4] □

While the previous proposition uses a nonconstructive argument avoiding to construct an explicit example with an exponential running time, it is possible to construct such an example even in the more restricted case of IS dynamics. To this end, it is possible to modify an example for MaxCut provided by Monien and Tscheuschner (2010) by essentially reverting the sequence of flips for MaxCut to obtain an execution of the IS dynamics. Thus, we generalize the previous proposition via a constructive proof.

Proposition 2. *The dynamics of IS deviations in symmetric ASHGs may require exponentially many rounds before converging to an IS partition.*

Proof. Let the agent set be $N = \{u_1, u_2, v_0\} \cup_{i=1}^n N_i$ with $N_i = \{v_i, u_{i,1}, u_{i,2}, u_{i,3}, u_{i,4}\}$, and consider the symmetric ASHG on this set of agents with utility values induced by the graph presented in Figure 3, where the weights of the building component G_i are depicted in Figure 4.[5] More precisely, the weight function is given by $f_n^i(k) = k + 5(2^{n-i+1} - 1)$.[6] All weights on missing edges are 0.

The underlying combinatorial structure consists of a short path G_0 together with n copies of the same graph with exponentially growing weights. Graph G_{i-1} and G_i are connected by an edge $\{v_{i-1}, u_{i,1}\}$.

Consider the partition of N indicated by the blue and green vertices and defined by $\pi = \{\{u_1, v_0\} \cup \bigcup_{i=1}^n \{v_i, u_{i,2}, u_{i,4}\}, \{u_2\} \cup \bigcup_{i=1}^n \{u_{i,1}, u_{i,3}\}\}$. We claim that there is an execution of the IS dynamics starting with π where agent v_i performs 2^{i+1} deviations for $i \in \{0, 1, \ldots, n\}$.

---

[4] We refer to the respective references for formal definitions of the involved combinatorial problems.

[5] Note that it is necessary in this example that the edge weights grow exponentially. If they were polynomially bounded, then the IS dynamics would run in polynomial time, because every deviation increases the social welfare.

[6] Note that there is a typo in the weight function by Monien and Tscheuschner (2010). They probably meant to use a similar weight function as the one used here.

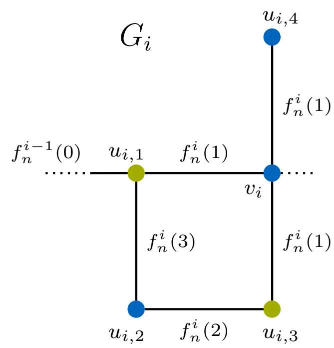


Figure 4: Weights of the building component G_i

We will recursively construct such a sequence of deviations. In the i-th step of the recursion, agent v_i will already perform 2^{i+1} deviations, and no agent in N_j will performs a deviation for $j > i$. Then, we will insert appropriate subsequences propagating through the graph. These insertions change the coalition $u_{i+1,1}$ was part of when v_i performs an IS deviation. However, this is not a problem, because the IS deviations of v_i are valid independently of the coalition that $u_{i+1,1}$ is part of. For $i = 0$, consider the sequence of deviations performed by (v_0, u_2, v_0), where v_0 performs $2 = 2^{0+1}$ deviations.

Now, let $k \geq 1$ and assume that the sequence is constructed for $k - 1$. We extend the sequence of deviations by inserting suitable subsequences. right before v_{k-1} performs her m-th deviation, then we insert

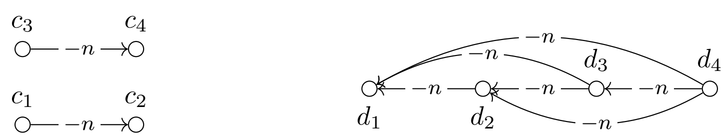
$$\begin{cases} (v_k, u_{k,3}, v_k, u_{k,2}, u_{k,3}, v_k, u_{k,1}) & \text{if } m \text{ odd} \\ (u_{k,2}, v_k, u_{k,1}) & \text{if } m \text{ even} \end{cases}$$

By the choice of the utility values and the initial partition, this sequence consists of NS deviations. Since all edge utility values are nonnegative, the sequence consists indeed of IS deviations. The most interesting deviations to check are the ones performed by agents $u_{i,1}$. Whenever they perform a deviation, they leave the coalition of $u_{i,2}$ and v_i to join the coalition of v_{i-1}. Indeed, this yields an improvement in utility, because $f_n^{i-1}(0) = 5(2^{n-i+2} - 1) > 4 + 10(2^{n-i+1} - 1) = f_n^i(3) + f_n^i(1)$. Note that after every even m, the subpartition of vertices in G_k is the same as in the initial partition π. Moreover, the agent v_k performs 2^{k+1} deviations.

In particular, for $k = n$, we have found an ASHG with a number of agents linear in n and (exponential) utility values which also require polynomial space. However, the constructed execution of the IS dynamics takes exponentially many rounds. □

5. Complexity of Stability under Majority Consent

In this section, we study stability under majority consent.



(a) AEG without MIS partition. (b) AEG without MOS partition.

Figure 5: Aversion-to-enemies games without MIS and MOS partitions from the proof ofProposition 3. Omitted edges have weight 1.

5.1. Aversion-To-Enemies Games

First, the existential results of Theorem 4 and Theorem 5 are contrasted with the non-existence of stable partitions in AEGs under the majority-based relaxations of the respective stability concepts.

Proposition 3. *There exists an AEG which contains no MIS (or MOS) partition.*

Proof. First, we provide an AEG with no MIS partition. Let $N = \{c_1, c_2, c_3, c_4\}$, i.e., there are $n = 4$ agents, and valuations defined as $v_{c_1}(c_2) = v_{c_3}(c_4) = -n$ and all other valuations set to 1. The AEG is illustrated in Figure 5a.

Assume for contradiction that there exists an MIS partition π. Then, $c_1 \notin \pi(c_2)$ and $c_3 \notin \pi(c_4)$. Also, $|\pi(c_1)| \leq 1$ (and $|\pi(c_3)| \leq 1$), because otherwise, c_2 (or c_4) would join via an MIS deviation. But then $\pi(c_1) = \{c_1\}$ and $\pi(c_3) = \{c_3\}$, and c_1 could deviate to join $\pi(c_3)$, a contradiction.

Second, we provide an AEG without MOS partition. Let $N = \{d_1, d_2, d_3, d_4\}$, and define valuations for all $i, j \in [4]$ with $i < j$ as $v_{d_i}(d_j) = 1$ and $v_{d_j}(d_i) = -4$. An illustration is provided in Figure 5b.

Assume for contradiction that there exists an MOS partition π. Then, every coalition $C \in \pi$ must fulfill $|C| \leq 2$. Otherwise, the agent of C with the second smallest index would form a singleton via an MOS deviation. In addition, there cannot be a singleton, because if some agent is in a singleton, there must be a second such agent, and then the one with the smaller index would join the other one. Hence, π consists of two pairs. But then d_1 would deviate to the pair not containing her, a contradiction. □

We can leverage the AEGs provided in the previous proposition as gadgets in reductions to show hardness of the associated decision problems. This can be interpreted as a more exact boundary (compared to Theorem 1) of the tractabilities encountered in Theorem 4 and Theorem 5 for the special case of AEGs.

Theorem 7. *It is* NP*-complete to decide if there exists an MIS (or MOS) partition in AEGs.*

5.2. Appreciation-Of-Friends Games

The utility restrictions in Theorem 7 are not as flexible as in the negative result for Nash stability in Theorem 1 or the positive results for unanimity-based dynamics in Theorem 4 and Theorem 5. In fact, the picture for majority-based notions is more diverse, because we obtain another positive result in the class of AFGs.

Theorem 8. *When starting from the grand coalition, the dynamics of MIS deviations converges after at most n rounds in AFGs.*

Proof. The key insight is that there can only be deviations to form a new singleton coalition yielding no more than n deviations. Let $\pi_0 = \{N\}$ be the initial partition, and consider a sequence of k MIS deviations

$$\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k.$$

We inductively define coalitions evolving from the grand coalition if removing the deviator as $G_0 = N$, and $G_j = G_{j-1} \setminus \{i_j\}$ for $j > 0$.

Now, we proceed to simultaneously prove the following claims by induction:

1. $\forall j \in [k] : \pi_{j-1}(i_j) = G_{j-1}$.
2. $\forall j \in [k] : \pi_j(i_j) = \{i_j\}$.
3. $\forall j \in [k] : \left\{i \in \pi_{j-1}(i_j) \colon v_{i_j}(i) = n\right\} = \emptyset$.

The base case $j = 1$ is immediate. For the induction step, let $2 \leq j \leq k$ and suppose the claims are true for all $1 \leq l < j$. We start with the first claim. By the induction hypothesis, $\pi_{j-1} = \{G_{j-1}\} \cup \{\{i_l\} \colon 1 \leq l < j\}$. This means that if $\pi_{j-1}(i_j) \neq G_{j-1}$, we must have $\pi_{j-1}(i_j) = \{i_j\}$, indicating $i_j = i_l$ for some $l < j$. Then, the welcoming coalition cannot be G_{j-1}, as i_j, by induction hypothesis, abandoned G_{l-1} due to not having any friends in G_{l-1}, and thus has, by $G_{j-1} \subseteq G_{l-1}$, no friends in G_{j-1}, either. The alternative is that i_j joins another singleton coalition $\{i_m\}$ to form a pair. However, if i_m abandoned G_m at some point $m < l$, then she dislikes i_j, and won't allow her to join. If i_m abandoned G_m at some point $m > l$, then i_j dislikes i_m, and has no incentive to join. Hence, $\pi_{j-1}(i_j) = G_{j-1}$. For the second claim, note that i_j cannot join another singleton $\{i_m\}$, because i_m abandoned G_{m-1} at some point $m < j$ and thus dislikes i_j. Hence, i_j must form a singleton $\pi_j(i_j) = \{i_j\}$, which she only wants to do if $\left\{i \in \pi_{j-1}(i_j) \colon v_{i_j}(i) = n\right\} = \emptyset$. This accomplishes the third claim, and completes the induction proof.

Finally, as there can be at most n singletons, the dynamics must terminate after at most n rounds. □

By contrast, we show in our next result that MOS partitions need not exist in AFGs. In other words, despite their conceptual complementarity, the stability concepts MOS and MIS lead to very different behavior in a natural subclass of ASHGs. The constructed game has a sparse friendship relation in the sense that almost all agents only have a single friend. After discussing the counterexample, we show how requiring slightly more sparsity yields a positive result.

Proposition 4. *There exists an AFG without an MOS partition.*

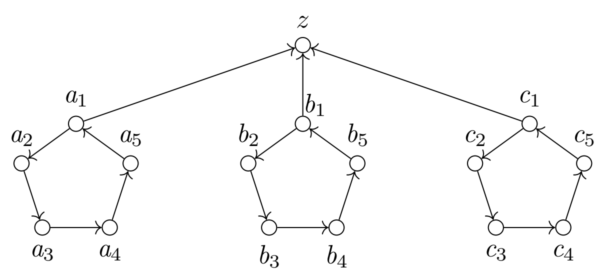


Figure 6: AFG without an MOS partition. The depicted (directed) edges represent friends, i.e., a utility of n, whereas missing edges represent a utility of -1.

Proof. We define the game formally. An illustration is given in Figure 6. Let $N = \{z\} \cup \bigcup_{x\in\{a,b,c\}} N_x$, where $N_x = \{x_i : i \in [5]\}$ for $x \in \{a, b, c\}$. In the whole proof, we read indices modulo 5, mapping to the respective representative in [5]. The utilities are given as:

- For all $i \in [5], x \in \{a, b, c\}$: $v_{x_i}(x_{i+1}) = n$.
- For all $x \in \{a, b, c\}$: $v_{x_1}(z) = n$.
- All other valuations are -1.

The AFG consists of 3 cycles with 5 agents each, together with a special agent that is liked by a fixed agent of each cycle and has no friends herself. The key insight to understanding why there exists no MOS partition is that agents of type x_1 where $x \in \{a, b, c\}$ have conflicting candidate coalitions in a potential MOS partition. Either, they want to be with z (a coalition that has to be small because z prefers to stay alone) or they want to be with x_2 which requires a rather large coalition containing their cycle.

Before going through the proof that this game has no MOS partition, it is instructional to verify that, for cycles of 5 agents, the unique MOS partition is the grand coalition, i.e., the unique MOS partition of the game restricted to N_x is $\{N_x\}$, where $x \in \{a, b, c\}$. This is a key idea of the construction and is implicitly shown in Case 2 of the proof for $x = b$.

Assume for contradiction that the defined AFG admits an MOS partition π. To derive a contradiction, we perform a case distinction over the coalition sizes of z.

<u>Case 1</u>: $|\pi(z)| = 1.$. In this case, it holds that $\pi(z) = \{z\}$. Then, $\pi(a_1) \in \{\{a_1, a_2\}, \{a_1, a_5\}\}$. Indeed, if $\pi(a_1) \neq \{a_1, a_2\}$, then a_1 has an NS deviation to join z, and is allowed to perform it unless $\pi(a_1) = \{a_1, a_5\}$. We may therefore assume that $\{a_i, a_{i+1}\} \in \pi$ for some $i \in \{1, 5\}$. Then, $\pi(a_{i-1}) = \{a_{i-1}, a_{i-2}\} =: C$. Otherwise, a_{i-1} can perform an MOS deviation to join $\{a_i, a_{i+1}\}$. But then a_{i+2} can perform an MOS deviation to join C. This is a contradiction and concludes the case that $|\pi(z)| = 1$.

Case 2: $|\pi(z)| > 1$.. Let $F := \{a_1, b_1, c_1\}$, i.e., the set of agents that have z as a friend. Note that z can perform an NS deviation to be a singleton. Hence, as π is MOS, $|F \cap \pi(v)| \geq |\pi(z)|/2$. In particular, there exists an $x \in \{a, b, c\}$ with $\pi(z) \cap N_x = \{x_1\}$. We may assume without loss of generality that $\pi(z) \cap N_a = \{a_1\}$. Then, $\pi(a_5) = \{a_4, a_5\}$. Otherwise, a_5 has an MOS deviation to join $\pi(z)$. Similarly, $\pi(a_3) = \{a_2, a_3\}$ (because of the potential deviation of a_3 who would like to join $\{a_4, a_5\}$). Now, note that $v_{a_1}(\{a_1, a_2, a_3\}) = n - 1$. We can conclude that $|\pi(z)| \leq 3$ as a_1 would join $\{a_2, a_3\}$ by an MOS deviation, otherwise. Hence, we find $x \in \{b, c\}$ with $N_x \cap \pi(z) = \emptyset$. Assume without loss of generality that $x = b$ has this property.

Assume first that $\pi(b_1) = \{b_1, b_5\}$. Then, $\pi(b_4) = \{b_3, b_4\}$. Otherwise, b_4 has an MOS deviation to join $\{b_1, b_5\}$. But then b_2 has an MOS deviation to join $\{b_3, b_4\}$, a contradiction. Hence, $\pi(b_1) \neq \{b_1, b_5\}$. Note that we have now excluded the only case where b_1 is not allowed to perform an NS deviation. In all other cases, no majority of agents prefers her to stay in the coalition. We can conclude that $b_2 \in \pi(b_1)$ because otherwise, b_1 can perform an MOS deviation to join $\pi(z)$. If $b_5 \notin \pi(b_1)$, then $\pi(b_5) = \{b_4, b_5\}$ (to prevent a potential deviation by b_5). But then b_3 has an MOS deviation to join them. Hence, $b_5 \in \pi(b_1)$. Similarly, if $b_4 \notin \pi(b_1)$, then $\pi(b_4) = \{b_3, b_4\}$ and b_2 has an MOS deviation to join $\{b_3, b_4\}$ (which is permissible because $b_5 \in \pi(b_1)$). Hence $\{b_1, b_2, b_4, b_5\} \subseteq \pi(b_1)$, and therefore even $N_b \subseteq \pi(b_1)$. Hence, b_1 has an MOS deviation to join $\pi(v)$ (recall that $|\pi(v)| \leq 3$). This is the final contradiction, and we can conclude that π is not MOS. $\square$

Note that most agents in the previous example have at most 1 friend (only three agents have 2 friends). By contrast, if every agent has at most one friend, MOS partitions are guaranteed to exist. This is interesting because it covers in particular directed cycles, which cause problems for Nash stability. The constructive proof of the following proposition can be directly converted into a polynomial-time algorithm.

Proposition 5. *Every AFG where every agent has at most one friend admits an MOS partition.*

Proof. We prove the statement by induction over n. Clearly, the grand coalition is MOS for $n = 1$. Now, assume that (N, v) is an AFG with $n \geq 2$ such that every agent has at most one friend. Consider the underlying directed graph $G = (N, A)$ where $(x, y) \in A$ if and only if $v_x(y) > 0$, i.e., y is a friend of x. By assumption, G has a maximum out-degree of 1, hence it can be decomposed into directed cycles and a directed acyclic graph.

Assume first that there exists $C \subseteq N$ such that C induces a directed cycle in G. We call an agent y *reachable* by agent x if there exists a directed path in G from x to y. Let $c \in C$ and define $R = \{x \in N : c \text{ reachable by } x\}$. Note that $C \subseteq R$ and that R is identical to the set of agents that can reach *any* agent in C. By induction, there exists an MOS partition π' of the subgame of (N, v) induced by $N \setminus R$ that is MOS. Define $\pi = \pi' \cup \{R\}$. We claim that π is MOS. Let $x \in N \setminus R$. By our assumptions on π', there exists no MOS deviation of x

to join $\pi(y)$ for $y \in N \setminus R$. In particular, if x is allowed to perform a deviation, then x must have a non-negative utility (otherwise, she can form a singleton coalition contradicting that π' is MOS). So her only potential deviations are to a coalition where she has a friend. Note that x has no friend in R. Indeed, if y was a friend of x in R, then c is reachable for x in G through the concatenation of (x, y) and the path from y to c. Hence, x has no MOS deviation. Now, let $x \in R$. Then, $v_x(\pi) > 0$ because she forms a coalition with her unique friend. By assumption, x has no friend in any other coalition. Therefore, x has no MOS deviation either.

We may therefore assume that G is a directed acyclic graph. Hence, there exists an agent $x \in N$ with in-degree 0. If x has no friend, let $T = \{x\}$. If x has a friend y, we claim that there exists an agent w such that (i) w is the friend of at least one agent and (ii) every agent that has w as a friend has in-degree 0, i.e., such agents are not the friend of any agent. We provide a simple linear-time algorithm that finds such an agent. We will maintain a tentative agent w that will continuously fulfill (i) and update w until this agent also fulfills (ii). Start with $w = y$. Note that this agent w fulfills (i) because y is a friend of x. If w is the friend of some agent z that is herself the friend of some other agent, update $w = z$. For the finiteness (and efficient computability) of this procedure, consider a topological order σ of the agents N in the directed acyclic graph G (Kahn, 1962), i.e., a function $\sigma\colon N \to [n]$ such that $\sigma(a) < \sigma(b)$ whenever $(a, b) \in A$. Note that if w is replaced by the agent z in the procedure, then $\sigma(z) < \sigma(w)$. Hence, w is replaced at most n times, and our procedure finds the desired agent w after a linear number of steps. Now, define $T = \{a \in N\colon w \text{ reachable by } a\}$, i.e., T contains precisely w and all agents that have w as a friend.

We are ready to find the MOS partition. By induction, we find a partition π' that is MOS for the subgame induced by $N \setminus T$. Consider $\pi = \pi' \cup \{T\}$. Then, $a \in T \setminus \{w\}$ has no incentive to deviate, because she has no friend in any other coalition and has w as a friend. Also, w is not allowed to perform a deviation, because the non-empty set of agents $T \setminus \{w\}$ unanimously prevents that. Possible deviations by agents in $N \setminus T$ can be excluded as in the first part of the proof because these agents have no friend in T. Together, we have completed the induction step and found an MOS partition. □

On the other hand, it is NP-complete to decide whether an AFG contains an MOS partition. For a proof, we use the game constructed in Proposition 4 as a gadget in a greater game. The difficulty is to preserve bad properties about the existence of MOS partitions because the larger game might allow for new possibilities to create coalitions with the agents in the counterexample.

Theorem 9. *Deciding whether an AFG contains an MOS partition is NP-complete.*

5.3. Friends-And-Enemies Games

We have already seen that friends-and-enemies games contain efficiently computable stable coalition structures with respect to the unanimity-based

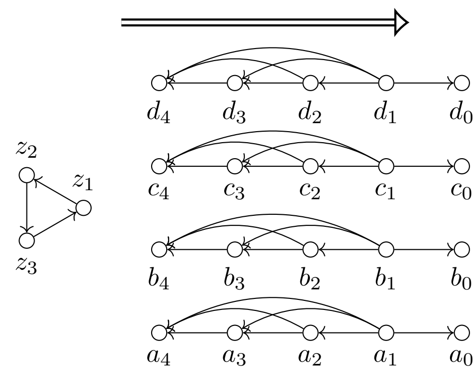


Figure 7: FEG without an MOS partition. The depicted (directed) edges represent friends. The double arrow means that every agent to the left of the tail of the arrow has every agent below the arrow as a friend.

stability concepts IS and CNS (cf. Corollary 1). In this section, we will see that the transition to majority-based consent crosses the boundary of tractability. The closeness to this boundary is also emphasized by the fact that it is surprisingly difficult to even construct No-instances for MOS and MIS, i.e., FEGs which do not contain an MOS or MIS partition, respectively. Indeed, the smallest such games that we can construct are games with 23 and 183 agents, respectively. We will start by considering MOS.

Proposition 6. *There exists an FEG without an MOS partition.*

Proof sketch. We only give a brief overview of the instance by means of the illustration in Figure 7. The FEG consists of a triangle of agents together with 4 sets of agents whose friendship relation is complete and transitive, together with one additional agent each that gives a temptation for the agent of the transitive substructures with the most friends.

An important reason for the non-existence of MOS partitions is that there is a high incentive for the transitive structures to form coalitions. This gives incentive to agents z_i to join them. If z_1, z_2, and z_3 are in disjoint coalitions, then they would chase each other according to their cyclic structure. If they are all in the same coalition, then agents x_0 for $x \in \{a, b, c, d\}$ prevent the complete transitive structures to be part of this coalition and other transitive structures are more attractive. □

In the previous proof, it is particularly useful to establish disjoint coalitions of groups of agents who dislike each other. On the other hand, if we make the further assumption that one agent from every pair of agents likes the other agent, then this does not work anymore and the grand coalition is MOS. This condition essentially means completeness of the friendship relation.[7] Note that

[7]Technically, the friendship relation may not be reflexive, but we can set $v_i(i) = 1$ for all

this proposition is not true for other stability concepts such as NS or even IS.

Proposition 7. *The grand coalition is MOS in every FEG with complete friendship relation.*

Proof. Let (N, v) be an FEG with complete friendship relation, and let π be the grand coalition. We claim that π is MOS. Suppose that there is an agent $x \in N$ who can perform an NS deviation to form a singleton.

Then, $v_x(N) < 0$ and therefore $|\{y \in N \setminus \{x\} \colon v_x(y) = -1\}| > \{y \in N \setminus \{x\} \colon v_x(y) = 1\}|$. Hence,

$$\begin{aligned}|F_{\text{in}}(N, x)| &\geq |\{y \in N \setminus \{x\} \colon v_x(y) = -1\}| \\ &> |\{y \in N \setminus \{x\} \colon v_x(y) = 1\}| \\ &\geq |F_{\text{out}}(N, x)|.\end{aligned}$$

In the first inequality, we use that x is a friend of all of her enemies. In the final inequality, we use that x can only be an enemy of her friends. Hence, x is not allowed to perform an MOS deviation. □

Still, the non-existence of MOS partitions in FEGs shown in Proposition 6 can be leveraged to prove an intractability result. Interestingly, in contrast to the proofs of Theorem 3 and Theorem 9, the next theorem merely uses the existence of an FEG without an MOS partition to design a gadget and does not exploit the specific structure of a known counterexample.

Theorem 10. *Deciding whether an FEG contains an MOS partition is NP-complete.*

In our next result, we construct an FEG without an MIS partition. Despite a lot of structure, the game is quite large encompassing 183 agents.

Proposition 8. *There exists an FEG without an MIS partition.*

Proof sketch. We illustrate the example with the aid of Figure 8 and briefly discuss some key features. Again, the central element is a directed cycle of three agents. These agents are connected to five copies of the same gadget. This gadget consists of a main clique $\{a_i^0, \dots, a_i^9\}$ of 10 mutual friends and further cliques that cause certain temptations for agents in the main clique. Cliques are linked by agents that have an incentive to be part of two cliques, which are part of disjoint coalitions. Since it is possible to balance all diametric temptations, the instance does not admit an MIS partition. □

Similar to Proposition 7, it is easy to see that the singleton partition is MIS in every FEG with complete enemy relation. Indeed, then an agent either has no incentive to join another agent, or the other agent will deny her consent. Hence, MIS can also prevent typical run-and-chase games which do not admit NS partitions. We are ready to prove hardness of deciding on the existence of MIS partitions in FEGs.

$i \in N$ in an FEG to formally achieve completeness.

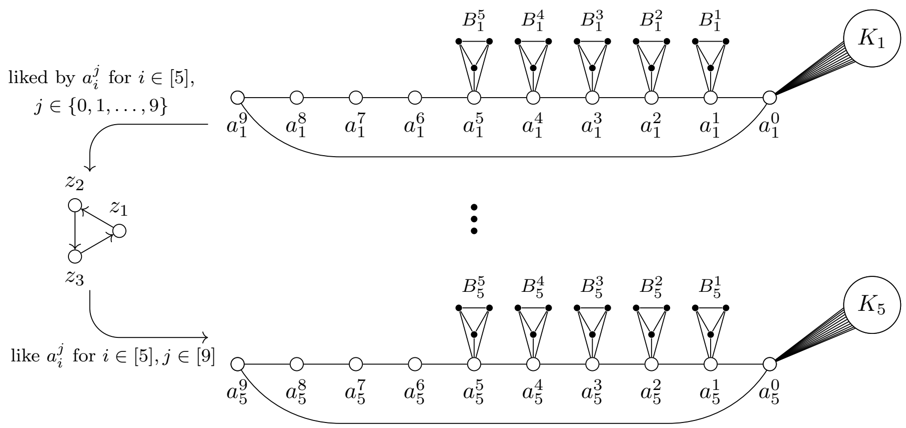


Figure 8: FEG without an MIS partition. The depicted edges represent friends. Undirected edges represent mutual friendship. For $i \in [5]$, some of the edges of agents in A_i are omitted. In fact, these agents form cliques. Also, each K_i represents a clique of 11 agents.

Theorem 11. *Deciding whether an FEG contains an MIS partition is NP-complete.*

5.4. Joint-Majority and Separate-Majorities Stability

The computational boundaries in this section encountered so far only hold for one-sided stability notions where either the welcoming or the abandoned coalition takes a vote. On the other hand, Theorem 6 shows that these are opposed by tractabilities under two-sided majority consent.

For general utilities, existence of SMS (and therefore JMS) partitions is not guaranteed anymore.

Proposition 9. *There exists an ASHG without SMS partition.*

Proof. Let $N = [5]$ and consider the utilities according to Table 1 below.

Table 1: Valuations for an ASHG without SMS partition.

v	1	2	3	4	5
1	0	2	−1	−3	1
2	1	0	2	−1	−3
3	−3	1	0	2	−1
4	−1	−3	1	0	2
5	2	−1	−3	1	0

See Figure 9 for a graphical representation of this example. We show that no partition can be SMS by an exhaustive case analysis. Let $+_{[5]}$ denote addition modulo 5, mapping to the representative in [5]. Assume for contradiction that π is SMS, and $C \in \pi$ is a coalition of largest cardinality.

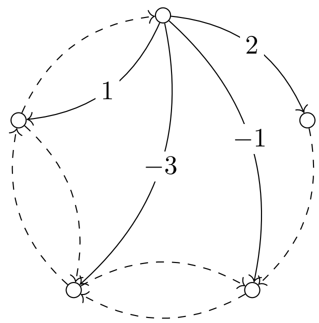


Figure 9: The ASHG without SMS partition from Proposition 9. Outgoing edges with weights have been drawn explicitly only for one agent, they are the same for each agent (up to rotation).

- Suppose $|C| = 5$. Then $\pi = \{N\}$, and all agents can form a singleton via an SMS deviation.
- Suppose $|C| = 4$. Then we can write it as $\{i, i +_{[5]} 1, i +_{[5]} 2, i +_{[5]} 3\}$ for some $i \in N$, and agent i can form a singleton via an SMS deviation.
- Suppose $|C| = 3$. Then it is either of the form $\{i, i +_{[5]} 1, i +_{[5]} 2\}$ or of the form $\{i, i +_{[5]} 1, i +_{[5]} 3\}$ for some $i \in N$. In the first case, agent $i +_{[5]} 2$ can form a singleton coalition, in the second case, agent $i +_{[5]} 3$ can form a singleton coalition.
- Suppose $|C| = 2$. Then π also has to contain a singleton $\{i\}$. If $\pi(i +_{[5]} 1) \in \{\{i +_{[5]} 1\}, \{i +_{[5]} 1, i +_{[5]} 2\}\}$, then i can join $i +_{[5]} 1$ via an SMS deviation. If $\pi(i +_{[5]} 1) \in \{\{i +_{[5]} 1, i +_{[5]} 3\}, \{i +_{[5]} 1, i +_{[5]} 4\}\}$, then $i +_{[5]} 1$ can join i via an SMS deviation.
- Suppose $|C| = 1$. Then any agent i can join $i +_{[5]} 1$ via an SMS deviation.

□

We can leverage game constructed in the proof of Proposition 9 to oppose Theorem 6 with a hardness result in general ASHGs.

Theorem 12. *Deciding whether an ASHG contains an SMS (or JMS) partition is* NP*-complete.*

6. Conclusion and Discussion

We studied stability based on single-agent deviations in additively separable hedonic games with a particular focus on games with restricted utility functions that can be naturally interpreted in terms of friends and enemies. We identified a computational boundary between Nash stability and stability with unanimous consent. The picture is less clear when deviations are governed by majority consent. While stable partitions always exist when considering both the abandoned and the welcoming coalition of the deviating agent, we obtain mostly negative results if only one of these coalitions is considered. Table 2 summarizes our results and compares them with related results from the literature. Notably, we obtain all of our positive results through the convergence of simple and natural dynamics. This also extends previously known results about IS. Aziz and Brandl (2012) obtain a polynomial algorithm essentially by running a dynamics from the singleton partition, whereas Dimitrov et al. (2006) take a different, graph-theoretical approach considering strongly connected components. The construction of CIS partitions by Aziz et al. (2013) is done by iteratively identifying specific coalitions, and it is not known whether CIS dynamics converge in polynomial time for natural starting partitions such as the singleton partition or grand coalition. An important tool in establishing our results concerning convergence of dynamics is the Deviation Lemma, a general combinatorial insight that allows us to study dynamics from a global perspective.

In addition, we have determined strong boundaries to the efficient computability of stable partitions. First, we resolve the computational complexity of computing CNS partitions, which considers the last open unanimity-based stability notion in unrestricted ASHGs. Second, our intractability concerning AFGs stands in contrast to the positive results for all other consent-based stability notions, and can also be circumvented by considering AFGs with a sparse friendship relation. Finally, we provide sophisticated hardness proofs for majority-based stability concepts in FEGs. These turn into computational feasibilities when transitioning to unanimity-based stability, or under further assumptions to the structure of the friendship graph.

A key step of all hardness results in restricted classes of ASHGs was to construct the first No-instances, that is, games that do not admit stable partitions for the respective stability notion. This is no trivial task as can be seen from the complexity of the constructed games. Once No-instances are found, we can leverage them as gadgets of hardness reductions, which is a typical approach for complexity results about hedonic games. We have provided both reductions where the explicit structure of the determined No-instances is used as well as reductions where the mere existence of No-instances is sufficient and used as a black box.

Together, our results give a complete picture of the computational complexity for all considered stability notions and game classes. Still, majority-based stability notions deserve further attention because they offer a natural degree of consent to perform deviations. Their thorough investigation in other classes of hedonic games might lead to interesting discoveries. Another intriguing further direction is to study further applications of the Deviation Lemma, particularly in domains other than coalition formation.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft under grants BR 2312/11-2 and BR 2312/12-1.

Table 2: Overview of our computational results. A red cell means existence of games without stable partition and usually comes with computational intractability. A green cell means that a stable partition can be constructed in polynomial time (Function-P), and in the case of our results even by executing a dynamics.
[a]: Aziz and Brandl (2012), [b]: Aziz et al. (2013), [c]: Dimitrov et al. (2006), [d]: Sung and Dimitrov (2010)

	General	*FEGs*	*AEGs*	*AFGs*
NS	NP-c[d]	NP-c (Th. 1)	NP-c (Th. 1)	NP-c (Th. 2)
IS	NP-c[d]	FP (Th. 4)	FP[a] (Th. 4)	FP[c] (Th. 4)
CNS	NP-c (Th. 3)	FP (Th. 5)	FP (Th. 5)	FP (Th. 5)
CIS	FP[b]	FP[b]	FP[b]	FP[b]
MIS	NP-c (Th. 7)	NP-c (Th. 11)	NP-c (Th. 7)	FP (Th. 8)
MOS	NP-c (Th. 7)	NP-c (Th. 10)	NP-c (Th. 7)	NP-c (Th. 9)
JMS	NP-c (Th. 12)	FP (Th. 6)	FP (Th. 6)	FP (Th. 6)
SMS	NP-c (Th. 12)	FP (Th. 6)	FP (Th. 6)	FP (Th. 6)

References

Abeledo, H., Rothblum, U.G., 1995. Paths to marriage stability. Discrete Applied Mathematics 63, 1–12.

Alcalde, J., Revilla, P., 2004. Researching with whom? Stability and manipulation. Journal of Mathematical Economics 40, 869–887.

Aziz, H., Brandl, F., 2012. Existence of stability in hedonic coalition formation games, in: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 763–770.

Aziz, H., Brandl, F., Brandt, F., Harrenstein, P., Olsen, M., Peters, D., 2019. Fractional hedonic games. ACM Transactions on Economics and Computation 7, 1–29.

Aziz, H., Brandt, F., Seedig, H.G., 2013. Computing desirable partitions in additively separable hedonic games. Artificial Intelligence 195, 316–334.

Aziz, H., Savani, R., 2016. Hedonic games, in: Brandt, F., Conitzer, V., Endriss, U., Lang, J., Procaccia, A.D. (Eds.), Handbook of Computational Social Choice. Cambridge University Press. chapter 15.

Ballester, C., 2004. NP-completeness in hedonic games. Games and Economic Behavior 49, 1–30.

Banerjee, S., Konishi, H., Sönmez, T., 2001. Core in a simple coalition formation game. Social Choice and Welfare 18, 135–153.

Bilò, V., Fanelli, A., Flammini, M., Monaco, G., Moscardelli, L., 2018. Nash stable outcomes in fractional hedonic games: Existence, efficiency and computation. Journal of Artificial Intelligence Research 62, 315–371.

Boehmer, N., Bullinger, M., Kerkmann, A.M., 2023. Causes of stability in dynamic coalition formation, in: Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI). Forthcoming.

Bogomolnaia, A., Jackson, M.O., 2002. The stability of hedonic coalition structures. Games and Economic Behavior 38, 201–230.

Brandt, F., Bullinger, M., 2020. Finding and recognizing popular coalition structures, in: Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 195–203.

Brandt, F., Bullinger, M., 2022. Finding and recognizing popular coalition structures. Journal of Artificial Intelligence Research 74, 569–626.

Brandt, F., Bullinger, M., Tappe, L., 2022. Single-agent dynamics in additively separable hedonic games, in: Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI), pp. 4867–4874.

Brandt, F., Bullinger, M., Wilczynski, A., 2021. Reaching individually stable coalition structures in hedonic games, in: Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI), pp. 5211–5218.

Brandt, F., Wilczynski, A., 2019. On the convergence of swap dynamics to Pareto-optimal matchings, in: Proceedings of the 15th International Conference on Web and Internet Economics (WINE), Springer-Verlag. pp. 100–113.

Bullinger, M., 2020. Pareto-optimality in cardinal hedonic games, in: Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 213–221.

Bullinger, M., 2022. Boundaries to single-agent stability in additively separable hedonic games, in: Proceedings of the 47th International Symposium on Mathematical Foundations of Computer Science (MFCS), pp. 26:1–26:15.

Bullinger, M., Kober, S., 2021. Loyalty in cardinal hedonic games, in: Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI), pp. 66–72.

Bullinger, M., Suksompong, W., 2023. Topological distance games, in: Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI). Forthcoming.

Carosi, R., Monaco, G., Moscardelli, L., 2019. Local core stability in simple symmetric fractional hedonic games, in: Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 574–582.

Cechlárová, K., Romero-Medina, A., 2001. Stability in coalition formation games. International Journal of Game Theory 29, 487–494.

Cseh, Á., 2017. Popular matchings, in: Endriss, U. (Ed.), Trends in Computational Social Choice. AI Access. chapter 6.

Dimitrov, D., Borm, P., Hendrickx, R., Sung, S.C., 2006. Simple priorities and core stability in hedonic games. Social Choice and Welfare 26, 421–433.

Dimitrov, D., Sung, S.C., 2007. On top responsiveness and strict core stability. Journal of Mathematical Economics 43, 130–134.

Drèze, J.H., Greenberg, J., 1980. Hedonic coalitions: Optimality and stability. Econometrica 48, 987–1003.

Elkind, E., Fanelli, A., Flammini, M., 2020. Price of pareto optimality in hedonic games. Artificial Intelligence 288, 103357.

Elkind, E., Wooldridge, M., 2009. Hedonic coalition nets, in: Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 417–424.

Fanelli, A., Monaco, G., Moscardelli, L., 2021. Relaxed core stability in fractional hedonic games, in: Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI), pp. 182–188.

Gairing, M., Savani, R., 2019. Computing stable outcomes in symmetric additively separable hedonic games. Mathematics of Operations Research 44, 1101–1121.

Gärdenfors, P., 1975. Match making: Assignments based on bilateral preferences. Behavioral Science 20, 166–173.

Hoefer, M., Vaz, D., Wagner, L., 2018. Dynamics in matching and coalition formation games with structural constraints. Artificial Intelligence 262, 222–247.

Kahn, A.B., 1962. Topological sorting of large networks. Communications of the ACM 5, 558–562.

Karp, R.M., 1972. Reducibility among combinatorial problems, in: Miller, R.E., Thatcher, J.W. (Eds.), Complexity of Computer Computations. Plenum Press, pp. 85–103.

Monien, B., Tscheuschner, T., 2010. On the power of nodes of degree four in the local max-cut problem, in: Proceedings of the 7th International Conference on Algorithms and Complexity (CIAC), Springer-Verlag. pp. 264–275.

Olsen, M., 2012. On defining and computing communities, in: Proceedings of the 18th Computing: Australasian Theory Symposium (CATS), pp. 97–102.

Roth, A.E., Vande Vate, J.H., 1990. Random paths to stability in two-sided matching. Econometrica 58, 1475–1480.

Saad, W., Han, Z., Basar, T., Debbah, M., Hjorungnes, A., 2011. Hedonic coalition formation for distributed task allocation among wireless agents. IEEE Transactions on Mobile Computing 10, 1327–1344.

Schäffer, A.A., Yannakakis, M., 1991. Simple local search problems that are hard to solve. SIAM Journal on Computing 20, 56–87.

Suksompong, W., 2015. Individual and group stability in neutral restrictions of hedonic games. Mathematical Social Sciences 78, 1–5.

Sung, S.C., Dimitrov, D., 2007. On myopic stability concepts for hedonic games. Theory and Decision 62, 31–45.

Sung, S.C., Dimitrov, D., 2010. Computational complexity in additive hedonic games. European Journal of Operational Research 203, 635–639.

Appendix

In the appendix, we provide the proofs missing in the body of the paper.

Appendix A. Missing Proofs in Section 3

Theorem 1. *Let $f^+ : \mathbb{N} \to \mathbb{Q}_{>0}$ and $f^- : \mathbb{N} \to \mathbb{Q}_{<0}$ be two polynomial-time computable functions satisfying $|f^-(m)| \geq f^+(m)$ for all $m \in \mathbb{N}$. Then, the problem of deciding whether an ASHG with utility values restricted to $\{f^-(n), f^+(n)\}$ has an NS partition is* NP*-complete.*

Proof. Let f^+, f^- be two functions as defined above and consider the class of ASHGs with utility values restricted to $\{f^-(n), f^+(n)\}$. We provide a reduction from the NP-complete problem Exact Cover by 3-Sets (E3C) (Karp, 1972). An instance of Exact Cover by 3-Sets consists of a tuple (R, S), where R is a ground set together with a set S of 3-element subsets of R. A 'yes'-instance is an instance such that there exists a subset $S' \subseteq S$ that partitions R. Given an instance (R, S) of E3C, for every $r \in R$, we define $S_r = \{s \in S : r \in s\}$, i.e., S_r comprises the elements of S containing r, and $n_r = |S_r|$.

Now, let (R, S) be an instance of E3C. We produce an ASHG (N, v) satisfying $v_i(j) \in \{f^-(n), f^+(n)\}$ for all $i, j \in N$ such that (R, S) has an exact cover if and only if (N, v) has an NS partition π. Define the agent set as $N = \{c\} \cup \bigcup_{s\in S} A^s \cup \bigcup_{r\in R}\{b_i^r : i \in [n_r - 1]\}$, where $A^s = \left\{a_{r_1}^s, a_{r_2}^s, a_{r_3}^s, a^s\right\}$ for $s = \{r_1, r_2, r_3\} \in S$. Hence, the agent set consists of copies of the elements in R corresponding to the frequency they occur in the sets of S minus 1, copies for the elements in sets of S together with one specific agent for each such set, and an auxiliary agent c. Now, define the following valuations v:

- For each $s \in S$, $a \neq a' \in A^s : v_a(a') = f^+(n)$.
- For each $r \in R, s \in S_r, i \in [n_r - 1] : v_{a_r^s}(b_i^r) = v_{b_i^r}(a_r^s) = v_{b_i^r}(c) = f^+(n)$.
- All other valuations are $f^-(n)$.

This reduction can be performed in polynomial time, as there are at most $4|S| + |R||S| + 1$ agents, and f^+, f^- can be computed in polynomial time. We claim that (R, S) admits an exact cover $S' \subseteq S$ if and only if (N, v) has an NS partition π.

$\Longrightarrow$. Suppose (R, S) has an exact cover $S' \subseteq S$. We construct an NS partition π.

- We have coalitions corresponding to the cover, i.e., for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \ldots, s_{n_r-1}$ and define for each $i \in [n_r - 1]$ the coalition $\{a_r^{s_i}, b_i^r\}$.

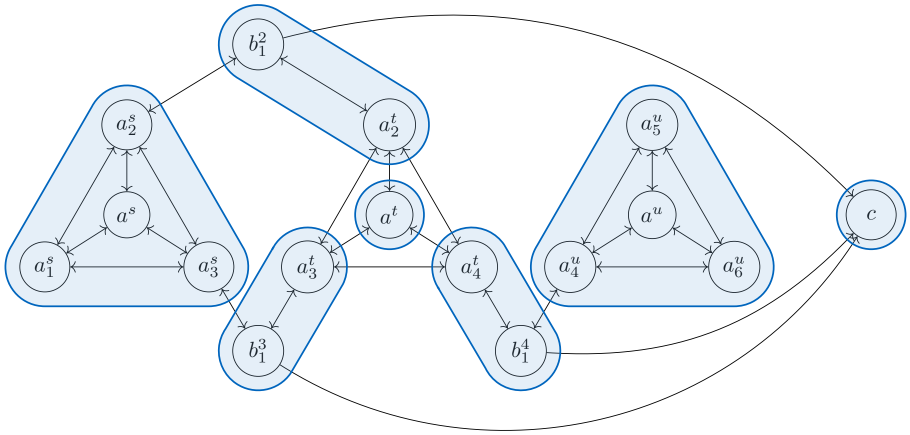


Figure A.10: The reduction from the proof of Theorem 1 for the Yes-instance of E3C $(\{1,\ldots,6\},\{s,t,u\})$ with $s=\{1,2,3\}, t=\{2,3,4\}$ and $u=\{4,5,6\}$. Drawn edges have weight $f^+(n)$, omitted edges have weight $f^-(n)$. The partition corresponding to the exact cover $\{s,u\}$ is highlighted.

- All agents a^s with $A^s \notin \pi$ are in a singleton: $\pi(a^s)=\{a^s\}$.
- Agent c is also in a singleton: $\pi(c)=\{c\}$.

To see that this partition is NS, we perform a case analysis for the various types of agents in order to show that no agent has an incentive to deviate.

- An agent a with $\pi(a)=A^s$ has $v_a(\pi)=3f^+(n)$, whereas every other coalition contains at most one agent she likes. So she has no incentive to deviate.
- An agent a_r^s with $\pi(a_r^s)\neq A^s$ is in a pair with an agent b_i^r, and so are the other two agents $a_{r'}^s$ from A^s. Thus, $v_{a_r^s}(\pi)=f^+(n)$, whereas every other coalition contains at most one agent she likes. So she has no incentive to deviate.
- An agent a^s with $\pi(a^s)\neq A^s$ is alone, but all other agents $a_r^s\in A^s$ are in a pair with an agent b_i^r that she dislikes, and as $f^+(n)+f^-(n)\leq 0$, she has no incentive to deviate.
- An agent b_i^r is in a pair with an agent a_r^s, so she has $v_{b_i^r}(\pi)=f^+(n)$. The best alternative would be joining c, which does not yield an improvement in utility, so she has no incentive to deviate.
- Finally, c has $v_c(\pi)=0$, which is her most desired outcome, as she dislikes all other agents.

Together, we conclude that π is NS.

$\Longleftarrow$. Suppose (N, v) contains an NS partition π. We show that then there must be an exact cover $S' \subseteq S$ of R. We begin with some observations:

1. Agent c must be in a singleton coalition, otherwise she would deviate to a singleton coalition.

2. Agents b_i^r must have utility $v_{b_i^r}(\pi) \geq f^+(n)$, otherwise they would join $\{c\}$.

3. Coalitions of agents a^s satisfy $\pi(a^s) \cap A^{s'} = \emptyset$ for $s' \neq s$. Suppose for contradiction that there is an agent $a \in \pi(a^s) \cap A^{s'}$. Consider the sets $A = \{i \in \pi(a^s) \colon v_a(i) = f^+(n)\}$ and $A' = \{i \in \pi(a^s) \colon v_{a^s}(i) = f^+(n)\}$. Then, we have $A \cap A' = \emptyset$. If $|A| \leq |A'|$, then a has an incentive to deviate to a singleton as she dislikes all agents from A' as well as a^s. Similarly, if $|A'| \leq |A|$, then a^s has an incentive to form a singleton coalition as she dislikes all agents from A as well as a.

4. Using observation 3, we must have $\pi(a^s) \neq \pi(b_i^r)$, as otherwise $v_{b_i^r}(\pi) \leq 0$, contradicting observation 2. Hence, we have $\pi(a^s) \subseteq A^s$ for all $s \in S$.

5. Now, consider an agent b_i^r. Define the sets $A = \{a_r^s \colon s \in S_r\}$ and $B = \left\{b_j^r \colon j \in [n_r - 1]\right\}$. By observation 2, we must have $|A \cap \pi(b_i^r)| \geq |\pi(b_i^r) \setminus A|$. We show that we must have $|A \cap \pi(b_i^r)| = |\pi(b_i^r) \setminus A|$. Suppose for contradiction that $|A \cap \pi(b_i^r)| > |\pi(b_i^r) \setminus A|$. Then, each agent $a_r^s \in A \cap \pi(b_i^r)$ has $v_{a_r^s}(\pi) \leq 0$ and would, by observation 4, rather deviate to $\pi(a^s)$. Moreover, we show that we must have $\pi(b_i^r) \setminus A \subseteq B$. Suppose for contradiction that this is not true. Then there are two cases. In the first case, there is an agent $b_j^{r'} \in \pi(b_i^r) \setminus A$ with $r \neq r'$. This agent dislikes all agents in A, and so would rather deviate to join $\{c\}$. In the second case, there is an agent $a_{r'}^s \in \pi(b_i^r) \setminus A$ with $r \neq r'$. This agent dislikes all but one agent from A as well as b_i^r, so would rather deviate to join $\pi(a^s)$.

Observation 5 shows that coalitions of agents b_i^r are of the form $A \uplus B$, where $A \subseteq \{a_r^s \colon s \in S_r\}, B \subseteq \left\{b_j^r \colon j \in [n_r - 1]\right\}$ and $|A| = |B|$. This leaves for each $r \in R$ exactly one agent a_r^s that is not in such a coalition. For these agents we have $\pi(a_r^s) = A^s$, yielding a cover $S' = \{s \in S \colon A^s \in \pi\}$. ☐

The proof of the next theorem is similar to the proof of Theorem 1. The essential difference is that we represent now every element in the ground set of an E3C-instance by a pair of agents.

Theorem 2. *Deciding whether an AFG has an NS partition is* NP*-complete.*

Proof. We provide another reduction from E3C. Let (R, S) be an instance of E3C. We produce an AFG (N, v) such that (R, S) has an exact cover if and only if (N, v) has a NS partition. Define the agent set $N = \{d\} \cup \bigcup_{s \in S} A^s \cup \bigcup_{r \in R} (\{c_1^r, c_2^r\} \cup \{b_i^r \colon i \in [n_r - 1]\})$, where $A^s = \{a_r^s \colon r \in s\}$ for $s \in S$.

Also, define the following valuations v:

- For each $s \in S, a \neq a' \in A^s : v_a(a') = n$.

- For each $r \in R, s \in S_r, i \in [n_r - 1] : v_{a_r^s}(b_i^r) = v_{b_i^r}(a_r^s) = v_{b_i^r}(d) = n$.
- For each $r \in R, s \in S_r : v_{c_1^r}(a_r^s) = v_{c_1^r}(c_2^r) = v_{c_2^r}(c_1^r) = v_{c_2^r}(d) = n$.
- All other valuations are -1.

This reduction can be performed in polynomial time, as there are only polynomially many agents. We now claim that (R, S) has an exact cover $S' \subseteq S$ if and only if (N, v) has a NS partition.

$\implies$. Suppose (R, S) has an exact cover $S' \subseteq S$. We construct a NS partition π.

- We have coalitions corresponding to the cover, i.e., for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \ldots, s_{n_r-1}$ and define for each $i \in [n_r - 1]$ the coalition $\{a_r^{s_i}, b_i^r\}$.
- For each $r \in R$ with $n_r > 1$, the agents c_1^r and c_2^r are in a pair $\{c_1^r, c_2^r\}$.
- Agent d is in a singleton $\{d\}$.

In this partition, each agent is together with some number of friends and no enemies. Every alternative coalition has at most as many friends as the current coalition, so no agent has incentive to deviate.

$\impliedby$. Suppose (N, v) has a NS partition π. We show that then there must be an exact cover $S' \subseteq S$ of R. We begin with some observations:

1. Agent d must be in a singleton coalition, because her value for any other agent is negative.
2. An agent c_2^r must be in a pair with c_1^r, otherwise she would join $\{d\}$.
3. An agent b_i^r must be in a coalition with at least one agent a_r^s, otherwise she would join $\{d\}$.
4. Agents a_r^s and $a_r^{s'}$ with $s \neq s'$ must be in distinct coalitions, otherwise c_1^r would join them.
5. Combining observations 3 and 4, we get that each agent b_i^r must be in a pair with exactly one agent a_r^s.

Define $S' = \{s \in S \colon \pi(a_r^s) \cap \{b_i^r \colon i \in [n_r - 1]\} = \emptyset \text{ for some } r \in s\}$. We claim that S' partitions R. First, we now know that for each $r \in R$, exactly $n_r - 1$ of the agents a_r^s must be in pairs with agents b_i^r. This leaves exactly one agent a_r^s not in a pair, and therefore not in a coalition with any agent from $\{b_i^r \colon i \in [n_r - 1]\}$. Hence, every agent in R is covered by S.

Now, assume for contradiction that there are $s_1, s_2 \in S$ with $s_1 \cap s_2 \neq \emptyset$. Let $j \in [2]$. Then, there exists $r_j \in s_j$ with $\pi(a^{s_j}_{r_j}) \cap \{b^{r_j}_i : i \in [n_r - 1]\} = \emptyset$ Since π is NS, it must be the case that $\pi(a^{s_j}_{r_j})$ contains at least one friend of $a^{s_j}_{r_j}$ and therefore $|\pi(a^{s_j}_{r_j}) \cap A^{s_j}| \geq 2$. Now, every agent in $A^{s_j} \setminus \pi(a^{s_j}_{r_j})$ can have at most one friend and would therefore perform an NS deviation to join $\pi(a^{s_j}_{r_j})$. Hence, there can be no such agent and therefore $A_{s_j} \subseteq \pi(a^{s_j}_{r_j})$. Hence, for $r \in s_1 \cap s_2$, at most $n_r - 2$ agents can be in pairs with agents b^r_i. This is a contradiction. Thus, the sets in S' are disjoint and therefore S' partitions R. □

Appendix B. Missing Proofs in Section 5.1

Theorem 7. *It is* NP*-complete to decide if there exists an MIS (or MOS) partition in AEGs.*

We split the proof into two separate reductions provided in Lemma 3 and Lemma 4. We start with the proof for MIS.

Lemma 3. *It is* NP*-complete to decide if there exists an MIS partition in AEGs.*

Proof. By reduction from E3C. Let (R, S) be an instance of E3C. We produce an AEG (N, v) such that (R, S) admits an exact cover if and only if (N, v) contains an MIS partition. Define $N = \bigcup_{s \in S} A^s \cup \bigcup_{r \in R} \bigcup_{i=1}^{n_r - 1} B^r_i$, where $A^s = \{a^s_{r_1}, a^s_{r_2}, a^s_{r_3}, a^s\}$ for $s = \{r_1, r_2, r_3\} \in S$, and $B^r_i = \{b^r_{i,j} : j \in [4]\}$ for $r \in R, i \in [n_r - 1]$. Define valuations v as:

- For each $s \in S, a \neq a' \in A^s$: $v_a(a') = 1$.
- For each $r \in R, s \in S_r, i \in [n_r - 1]$: $v_{a^s_r}(b^r_{i,1}) = v_{b^r_{i,1}}(a^s_r) = 1$.
- Each B^r_i has internal valuations as in the first example of Proposition 3, i.e., if v' denotes the valuations of this example, then $v_{b^r_{i,j}}(b^r_{i,k}) = v'_{c_j}(c_k)$, where the negative valuations are adapted to the specific number of agents in the instance.
- All other valuations are $-n$.

We proceed to prove correctness of the reduction.

$\Longrightarrow$. Suppose (R, S) has an exact cover $S' \subseteq S$. We construct an MIS partition π as follows.

- We have coalitions corresponding to the cover, i.e. for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \ldots, s_{n_r - 1}$ and define for each $i \in [n_r - 1]$ the coalitions $\{a^{s_i}\}$, $\{a^{s_i}_r, b^r_{i,1}\}$, $\{b^r_{i,2}, b^r_{i,4}\}$, and $\{b^r_{i,3}\}$.

No agent has an incentive to deviate, making the partition NS and thus MIS.

$\Longleftarrow$. Suppose (N, v) has an MIS partition π. We construct an exact cover $S' \subseteq S$. We begin with some observations:

1. No agent is in a coalition with someone she dislikes, otherwise she would deviate to a singleton coalition. In particular, this means $\pi(a^s) \subseteq A^s$ and $\pi(b^r_{i,j}) \subseteq B^r_i$ for $j \in \{2, 3, 4\}$.

2. Each agent of type $b^r_{i,1}$ must be in a coalition with exactly one agent a^s_r. If $\pi(b^r_{i,1}) \subseteq B^r_i$, we would contradict the fact that the subgame induced by B^r_i has no stable partition (see Proposition 3). As $b^r_{i,1}$ cannot form a coalition with someone she dislikes, at least one agent c of the type a^s_r must be in her coalition. Finally, no other agent giving positive utility to $b^r_{i,1}$ can be in a common coalition with c.

Now, we know that for each $r \in R$, exactly $n_r - 1$ of the agents a^s_r must be in pairs with $b^r_{i,1}$. This leaves exactly one agent a^s_r not in a pair. We claim that for these agents we have $\pi(a^s_r) = A^s$. Indeed, it is clear that we then must have $\pi(a^s_r) \subseteq A^s$. If $\pi(a^s_r) = \{a^s_r\}$, she would deviate to join $\pi(a^s)$. Then, $|\pi(a^s_r)| \geq 2$, and members from $A^s \setminus \pi(a^s_r)$ would have an incentive to join $\pi(a^s_r)$. It follows that $A^s \setminus \pi(a^s_r) = \emptyset$, and therefore $\pi(a^s_r) = A^s$. Hence, we obtain a cover $S' = \{s \in S \colon A^s \in \pi\}$. □

Note that it can be shown that a partition in the reduced instances in the reduction of the previous lemma is NS if and only if it is MIS. Hence, the lemma provides yet another proof to the respective statement about Nash stability first shown by Sung and Dimitrov (2010) (and already revisited in Theorem 1). We proceed with the complementary proof for MOS.

Lemma 4. *It is* NP*-complete to decide if there exists an MOS partition in AEGs.*

Proof. Again, we reduce from E3C. Let (R, S) be an instance of E3C. We produce an AEG (N, v) with agent set $N = \bigcup_{s \in S} A^s \cup \bigcup_{r \in R} \bigcup_{i=1}^{n_r - 1} B^r_i$, where $A^s = \left\{a^s_{r_1}, a^s_{r_2}, a^s_{r_3}, a^s\right\}$ for $s = \{r_1, r_2, r_3\} \in S$, and $B^r_i = \left\{b^r_{i,j} \colon j \in [4]\right\}$ for $r \in R, i \in [n_r - 1]$. Define the following valuations v:

- For each $s \in S, a \neq a' \in A^s$: $v_a(a') = 1$.
- For each $r \in R, s \in S_r, i \in [n_r - 1]$: $v_{a^s_r}(b^r_{i,1}) = 1$.
- Each B^r_i has internal valuations as in the second example constructed in the proof of Proposition 3, i.e., if v' are the valuations from this example, then $v_{b^r_{i,j}}(b^r_{i,k}) = v'_{d_j}(d_k)$, where the negative valuations are adapted to the specific number of agents in the instance.
- All other valuations are $-n$.

We claim that (R, S) has an exact cover if and only if (N, v) has an MOS partition.

$\Longrightarrow$. Suppose (R, S) has an exact cover $S' \subseteq S$. We construct an MOS partition π.

- We have coalitions corresponding to the cover, i.e. for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \dots, s_{n_r-1}$ and define for each $i \in [n_r - 1]$ the coalitions $\{a^{s_i}\}$, $\left\{a_r^{s_i}, b_{i,1}^r\right\}$, $\left\{b_{i,2}^r, b_{i,3}^r\right\}$, and $\left\{b_{i,4}^r\right\}$.

The only agents that have an incentive to deviate are agents of types $b_{i,1}^r$ and $b_{i,3}^r$. However, there is some $s \in S$ such that $\pi(b_{i,1}^r) = \left\{b_{i,1}^r, a_r^s\right\}$, and a_r^s ensures that b_i^r cannot leave. Similarly, $\pi(b_{i,3}^r) = \left\{b_{i,3}^r, b_{i,2}^r\right\}$, and $b_{i,2}^r$ ensures that $b_{i,3}^r$ cannot leave. Note that agents a^s for $s \notin S'$ cannot deviate, because all their friends form a coalition with an enemy. Hence, π is MOS.

$\Longleftarrow$. Suppose now that (N, v) has an MOS partition π. We construct an exact cover $S' \subseteq S$. First, we make some observations:

1. Agents $b_{i,2}^r$ must have $\pi(b_{i,2}^r) \subseteq B_i^r$. If there was an agent $a \in \pi(b_{i,2}^r) \setminus B_i^r$, then, as $v_{b_{i,2}^r}(a) = -n$, $b_{i,2}^r$ would rather be in a singleton, and could form one, as $\left|F_{\text{out}}(\pi(b_{i,2}^r), b_{i,2}^r)\right| \geq |\{a\}| = 1 = \left|\left\{b_{i,1}^r\right\}\right| \geq \left|F_{\text{in}}(\pi(b_{i,2}^r), b_{i,2}^r)\right|$.
2. Using observation 1, we can conclude that agents $b_{i,3}^r$ must also have $\pi(b_{i,3}^r) \subseteq B_i^r$.
3. Using observations 1 and 2, we can conclude that agents $b_{i,4}^r$ must also have $\pi(b_{i,4}^r) \subseteq B_i^r$.
4. Agents $a \in A^s$ and $a' \in A^{s'}$ with $s \neq s'$ satisfy $\pi(a) \neq \pi(a')$. For contradiction, suppose this is not the case, i.e., there are $a \in A^s$ and $a' \in A^{s'}$ with $s \neq s'$ such that $\pi(a) \neq \pi(a') =: C$. Clearly, both prefer to be in a singleton coalition. Further, we can assume without loss of generality that $|A^s \cap C| \leq \left|A^{s'} \cap C\right|$ (otherwise, we can just swap them). Then, as $|F_{\text{out}}(C, a)| \geq \left|A^{s'} \cap C\right| \geq |A^s \cap C| > |F_{\text{in}}(C, a)|$, agent a could deviate to form a singleton coalition, a contradiction.
5. Agents $b_{i,1}^r$ must be in a coalition with no other agents from B_i^r and at least one other agent from $N \setminus B_i^r$. This follows from observations 1, 2, and 3 in conjunction with the fact that the subgame induced by B_i^r is identical to the example from the second part of Proposition 3 which has no MOS partition. Due to the valuations for agent $b_{i,1}$, some agent a_r^s must be in her coalition, and due to observation 4, there can be at most one such agent in her coalition. If there were further agents from A^s in her coalition, $b_{i,1}^r$ could deviate to a singleton coalition. Thus, the only possibility is that $b_{i,1}^r$ is in a pair with exactly one agent a_r^s.

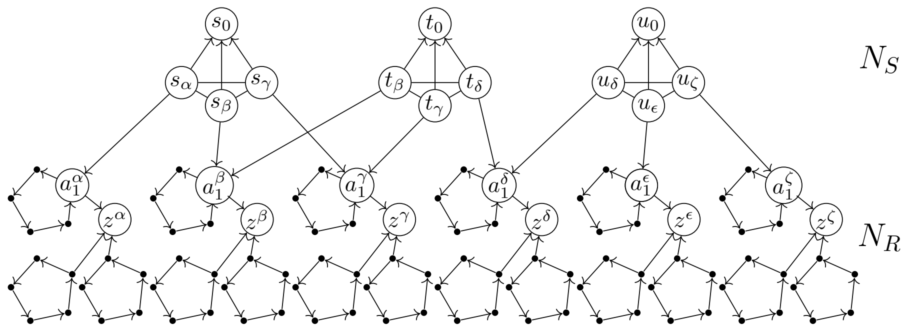


Figure C.11: Schematic of the reduction from the proof of Theorem 9. We depict the reduced instance for the instance (R, S) of E3C where $R = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta\}$ and $S = \{s, t, u\}$ with $s = \{\alpha, \beta, \gamma\}$, $t = \{\beta, \gamma, \delta\}$, and $u = \{\delta, \epsilon, \zeta\}$. Directed edges indicate a utility of n, and missing edges a utility of -1. Every element in R is represented by a gadget identical to the game in Proposition 4.

We now know that for each $r \in R$, exactly $n_r - 1$ of the agents a_r^s must be in pairs with $b_{i,1}^r$. This leaves exactly one agent a_r^s not in a pair. For these agents we have $\pi(a_r^s) \subseteq A^s$. Also, $\pi(a^s) \subseteq A^s$, as any agent outside would like to leave and there is at most 1 vote for her to stay. Consequently, $|\pi(a_r^s)| \geq 2$, and members from $A^s \setminus \pi(a_r^s)$ would have an incentive to join $\pi(a_r^s)$. Hence, $\pi(a_r^s) = A^s$, and we obtain a cover $S' = \{s \in S \colon A^s \in \pi\}$. □

Appendix C. Missing Proof in Section 5.2

Theorem 9. *Deciding whether an AFG contains an MOS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let (R, S) be an instance of E3C. We define an ASHG (N, v) as follows. Let $N = N_R \cup N_S$ where $N_R = \cup_{r \in R} N^r$ and $N_S = \cup_{s \in S} N_s$ with $N^r = \{a_i^r, b_i^r, c_i^r \colon i \in [5]\} \cup \{z^r\}$ for $r \in R$ and $N_s = \{s_r \colon r \in s\} \cup \{s_0\}$ for $s \in S$. In the whole proof, we read indices of agents a_i^r, b_i^r, and c_i^r modulo 5, mapping to the representative in [5].

We define utilities v as follows:

- For all $s \in S$, $r \in s$: $v_{s_r}(s_0) = n$.
- For all $s \in S$, $r, r' \in s$, $r \neq r'$: $v_{s_r}(s_{r'}) = n$.
- For all $s \in S$, $r \in s$: $v_{s_r}(a_1^r) = n$.
- For all $r \in R$, $i \in [5]$, and $x \in \{a, b, c\}$: $v_{x_i}^r(x_{i+1}^r) = n$.
- For all $r \in R$, $x \in \{a, b, c\}$: $v_{x_1^r}(z) = n$.
- All other valuations are -1.

An illustration of the reduction is provided in Figure C.11. Intuitively, the reduced instance consists of two types of gadgets. The elements in the ground set R are represented by R-gadgets which are subgames identical to the counterexample in Proposition 4. The sets in S are represented by S-gadgets consisting of a triple of agents representing its elements in R which are linked to the respective R-gadgets. Furthermore, there is one special agent without any friends attracting the other agents in the S-gadget.

We claim that (R, S) is a Yes-instance if and only if the reduced AFG contains an MOS partition.

$\Longrightarrow$. Suppose first that $S' \subseteq S$ partitions R. We define a partition π by taking the union of the following coalitions:

- For $r \in R$, $x \in \{a, b, c\}$, form $\{x_2^r, x_3^r\}$, $\{x_4^r, x_5^r\}$, and $\{b_1^r, c_1^r, z^r\}$.
- For $s \in S', r \in s$, form $\{s_r, a_1^r\}$.
- For $s \in S'$, form $\{s_0\}$.
- For $s \in S \setminus S'$, form N_s.

We prove that π is MOS by performing a case analysis to show that no agent can perform a deviation.

- For $r \in R$ and $x \in \{a, b, c\}$, the agents x_3^r and x_5^r are not allowed to perform an MOS deviation. Moreover, the agents x_2^r and x_4^r are in their most preferred coalitions, and have therefore no incentive to perform a deviation.
- For $r \in R$, the agents a_1^r and z^r are not allowed to perform an MOS deviation.
- For $r \in R$ and $x \in \{b, c\}$, the agent x_1^r has no incentive to deviate. It holds that $v_{x_1^r}(\pi) = n - 1$, whereas no deviation increases her utility. In particular, joining $\pi(x_2^r)$ only yields the same utility.
- For $s \in S$ and $r \in s$, the agent s_r has at most one friend after any possible deviation. However, she has at least two friends in π, and therefore no incentive to perform a deviation.
- For $s \in S'$, the agent s_0 is in her most preferred coalition and has no incentive to perform a deviation. Finally, for $s \in S \setminus S'$, the agent s_0 is not allowed to perform an MOS deviation.

$\Longleftarrow$. Conversely, assume that the reduced instance contains an MOS partition π. We show that it originates from a Yes-instance. We split the proof into several claims.

Claim 1. *For all $s \in S$, it holds that $\pi(s_0) = \{s_0\}$ or $N_s \subseteq \pi(s_0)$.*

Proof. Let $s \in S$, say $s = \{u, w, x\}$, and define $C = \pi(s_0)$ and $D = \{s_u, s_w, s_x\}$. Assume that $C \supsetneq \{s_0\}$. Then, since s_0 has no friends, she would prefer to stay in a singleton coalition. Hence, $C \cap D \neq \emptyset$, say $s_u \in C$.

Assume for contradiction that $D \setminus C \neq \emptyset$, say $s_w \notin C$. Then, $s_x \in \pi(s_w)$. Indeed, if $s_x \notin \pi(s_w)$, then s_w has at most one friend in her coalition, and no agent would prevent her from performing an MOS deviation to join C. Hence, $s_x \in \pi(s_w)$. Then, $C = \{s_0, s_u\}$, as s_0 could leave her coalition to form a singleton coalition if any other agent was part of it. But then, s_u has an incentive to join $\pi(s_w)$, and could perform a valid MOS deviation to do so. This is a contradiction and therefore $D \subseteq C$. $\lhd$

In the next claim, we improve upon Claim 1 and show that there are in fact only two possible coalitions for s_0.

Claim 2. *For all $s \in S$, it holds that $\pi(s_0) = \{s_0\}$ or $\pi(s_0) = N_s$.*

Proof. Let $s \in S$ and define $C = \pi(s_0)$. Assume that $C \supsetneq \{s_0\}$. By Claim 1, it holds that $N_s \subseteq C$ and since s_0 has an NS deviation to form a singleton coalition, even $|C| \leq 6$. This means in particular that every agent $y \in C \setminus N_s$ must have a friend in C. Indeed, if this was not the case, then such an agent y would like to deviate to form a singleton coalition and this is an MOS deviation as it is supported by at least three agents in N_s. Hence, $C \setminus N_s \neq \emptyset$ can only happen if there are two more agents in C who are a friend of each other. By the design of the utilities, the only possibility for this to happen is that there exists $t \in S$ with $t \neq s$ and $u, v \in t$ with $C = N_s \cup \{t_u, t_v\}$. Then, by Claim 1, $\{t_0\} \in \pi$, implying that t_u has an MOS deviation to join t_0. This is a contradiction and we can therefore conclude that $\pi(s_0) = N_s$. $\lhd$

Next, we consider the coalitions of other agents in gadgets related to sets in S.

Claim 3. *For all $s \in S$ and $r \in R$, it holds that $\pi(s_r) = \{s_r, a_1^r\}$ or $N_s \setminus \{s_0\} \subseteq \pi(s_r)$.*

Proof. Let $s \in S$, say $s = \{r, u, w\}$, and define $C = \pi(s_r)$. If $s_0 \in C$, then $C = N_s$ by Claim 2 and the assertion is true. Suppose therefore that $s_0 \notin C$. Assume now that there is $x \in s$ with $s_x \notin C$, say $s_u \notin C$. If $s_w \notin C$, then no agent in C has s_r as a friend and could therefore vote against a deviation. Moreover, since the deviation of s_r to join s_0 is not an MOS deviation, it must be the case that $v_{s_r}(\pi) = n$, which can, under the given assumptions, only be the case if $\pi(s_r) = \{s_r, a_1^r\}$.

It remains to consider the case that $s_w \in C$. But then, s_u is in a coalition with at most one friend (note that it is excluded that $s_0 \in \pi(s_u)$ by Claim 2) and no agent in her coalition has her as a friend. Hence, s_u has an MOS deviation to join C, a contradiction. Together, we have shown that if there is $x \in s$ with $s_x \notin C$, then $\pi(s_r) = \{s_r, a_1^r\}$, which proves this claim. $\lhd$

In the next claim, we gain even more insight on the coalitions of agents of the type s_r.

Claim 4. *For all $s \in S$, $r \in s$, and $u \in R$, it holds that if $\pi(s_r) \cap N^u \neq \emptyset$, then $r = u$ and $\pi(s_r) = \{s_r, a_1^u\}$.*

Proof. Let $s \in S$, $r \in s$, and $u \in R$. The assertion is true if $\pi(s_r) = \{s_r, a_1^r\}$. Hence, by Claim 3, we may assume that $N_s \setminus \{s_0\} \subseteq C$. we will show that $\pi(s_r) \cap N^u = \emptyset$. First, note that since z^u has an NS deviation to form a singleton coalition whenever she is not in such a coalition already and because only three agents have z^u as a friend, it holds that z^u forms a coalition with at most two agents that have her as an enemy. This implies in particular that $z^u \notin C$ and that $|\pi(z^u)| \leq 6$.

Assume for contradiction that there exists an agent $y \in N^u \cap C$. We already know that $y \neq z^u$. Next, if $y \neq a_1^u$, then y must have a friend in C. Indeed, at most one agent in C can have y as a friend, but the three agents in $N_s \setminus \{s_0\}$ favor y to leave. Hence, y could perform an MOS deviation to form a singleton coalition, otherwise. In addition, if $y = a_1^u$, then y must also have a friend in C. Note that at most two agents in $(N^u \cup N_s) \cap C$ favor her to stay while all other agents in $(N^u \cup N_s) \cap C$ (of which there are at least 2 agents) favor her to leave. The only possibility that there is another agent who favors a_1^u to stay is if there exists $t \in S$ with $u \in t$ and $t_u \in C$. But then, Claim 3 implies that $N_t \setminus \{t_0\} \subseteq C$, a majority of which favors a_1^u to leave. Together, a_1^u is favored to leave C by a (weak) majority of agents. Therefore, she must not have an incentive to form a singleton coalition, and therefore has a friend in C.

Now, assume that there exists $x \in \{a, b, c\}$ and $i \in [5]$ with $x_i^u \in C$. Then, our previous observation implies that $\{x_i^u : i \in [5]\} \subseteq C$. Hence, $|C| \geq 8$ and therefore $v_{x_1^u}(\pi) \leq n-6 < n-5 \leq v_{x_1^u}(\pi(z^u) \cup \{x_1^u\})$. Hence, x_1^u could perform an MOS deviation, a contradiction. Therefore, we have shown that $\pi(s_r) \cap N^u = \emptyset$. ◁

Now, we show that coalitions of agents in different sets of the type N^r are disjoint.

Claim 5. *For all $r, u \in R$ and agents $w \in N^r, y \in N^u$, it holds that $\pi(w) \cap \pi(y) = \emptyset$.*

Proof. Let $r, u \in R$ and assume for contradiction that there exist agents $w \in N^r$ and $y \in N^u$ with $\pi(w) = \pi(y)$. Define $C = \pi(w)$. By Claim 2 and Claim 4, it holds that $C \cap N_s = \emptyset$ for all $s \in S$. We may assume without loss of generality that $|C \cap N^r| \leq |C \cap N^u|$. Since every agent in $C \cap N^r$ is preferred to leave by a majority of agents in C, it holds that $z^r \notin C$ and every agent in $C \cap N^r$ must have a friend in C. The remaining proof of this step is similar to the proof of Claim 4. Let $x \in \{a, b, c\}$ and $i \in [5]$ with $x_i^r \in C$. Then, $\{x_i^r : i \in [5]\} \subseteq C$ and therefore $|C| \geq 10$. As in the previous claim, $|\pi(z^r)| \leq 6$. Hence, $v_{x_1^r}(\pi) \leq n-8 < n-5 \leq v_{x_1^r}(\pi(z^r) \cup \{x_1^r\})$, a contradiction. ◁

Finally, we can conclude the proof by showing that there exists $S' \subseteq S$ partitioning R. Therefore, let $S' = \{s \in S : \pi(s_r) = \{s_r, a_1^r\} \text{ for some } r \in s\}$. We show that S' partitions R by showing that it covers all elements from R and that its elements are disjoint sets.

For the first part, let $r \in R$. By the proof of Proposition 4, if $\pi(y) \subseteq N^r$ for all $y \in N^r$, then the partition π is not MOS. Hence, some agent in N^r must form a coalition with an agent outside of N^r. Combining Claim 2, Claim 4, and Claim 5, this can only be the case if there exists $s \in S$ with $r \in s$ and $\pi(s_r) = \{s_r, a_1^r\}$. Consequently, S' covers R.

For the second part, assume for contradiction that some element in R is covered at least twice by sets in S'. Then, there exists $s \in S'$ with $r \in s$ and $\{s_r, a_1^r\} \notin \pi$. By Claim 3, $N_s \setminus \{s_0\} \subseteq \pi(s_r)$. But then, according to the definition of S', it follows that $s \notin S'$, a contradiction. Hence, the elements of S' are disjoint sets. This completes the proof. □

Appendix D. Missing Proofs in Section 5.3

In this section, we provide missing proofs about majority-based stability concepts in FEGs.

Proposition 6. *There exists an FEG without an MOS partition.*

Proof. Recall that the game is illustrated in Figure 7. Formally, let $N = N_z \cup N_a \cup N_b \cup N_c \cup N_d$, where $N_z = \{z_1, z_2, z_3\}$ and $N_x = \{x_0, x_1, x_2, x_3, x_4\}$ for $x \in \{a, b, c, d\}$. Utilities are given as

- $v_x(y) = 1$ if $(x, y) \in \{(z_1, z_2), (z_2, z_3), (z_3, z_1)\}$,
- $v_{x_i}(x_j) = 1$ if $x \in \{a, b, c, d\}, i, j \in [4], i < j$,
- $v_{x_1}(x_0) = 1$ if $x \in \{a, b, c, d\}$,
- $v_{z_i}(x_j) = 1$ if $x \in \{a, b, c, d\}, i \in [3], j \in [4]$, and
- all other valuations are -1.

Assume for contradiction that this FEG admits an MOS partition π. We will derive a contradiction in 4 steps. First, Claim 6 describes possible coalitions of agents x_0 where $x \in \{a, b, c, d\}$. Second, Claim 7 establishes that coalitions from agents of different sets of N_x, $x \in \{a, b, c, d\}$, are disjoint. Then, Claim 8 excludes that all agents in N_z are in a joint coalition. Finally, we complete the proof by performing a case analysis for two disjoined coalitions containing different agents from N_z.

Claim 6. *It holds that $\pi(x_0) \in \{\{x_0\}, \{x_0, x_1\}\}$ for $x \in \{a, b, c, d\}$.*

Proof. Let $x \in \{a, b, c, d\}$ and suppose that $|\pi(x_0)| > 1$. Then, x_0 has an NS deviation to form a singleton. The claim follows because the only agent that prevents her to leave the coalition is x_1. ◁

Claim 7. *It holds that $x_i \notin \pi(y_j)$ for $x, y \in \{a, b, c, d\}, x \neq y$, and $i, j \in [4]$.*

Proof. Assume for contradiction that there exist $x, y \in \{a, b, c, d\}, x \neq y$, and $i, j \in [4]$ with $x_i \in \pi(y_j)$. Without loss of generality, $x = a$ and $y = b$. Define $\Gamma := \pi(b_j)$. Again, without loss of generality, we may assume that $|\Gamma \cap N_a| \geq |\Gamma \cap N_b|$. Let $j^* = \min\{j \in [4] \colon b_j \in \Gamma\}$.

By Claim 6, $x_0 \notin \Gamma$ for $x \in \{a, b, c, d\}$. Hence, b_{j^*} wants to perform an NS deviation to form a singleton and is only favored to stay by agents in N_z. As $a_i \in F_{\text{out}}(\Gamma, b_{j^*})$, at least two agents must favor b_{j^*} to stay. We conclude that

- $|\Gamma \cap N_z| \geq 2 \quad (*)$
- $|\Gamma \setminus N_z| \leq 3 \quad (**)$

There, $(**)$ follows because at most 3 agents favor b_{j^*} to stay, and she can therefore have at most two enemies. To conclude this step, we distinguish two cases.

Case 1: It holds that $|N_z \cap \Gamma| = 3$, i.e., $N_z \subseteq \Gamma$.. We consider now the agents in N_c. By Claim 6, $(*)$, and $N_z \subseteq \Gamma$, we derive that $\pi(c_i) \subseteq N_c \setminus \{c_0\}$ for $i = 2, 3, 4$, and $\pi(c_1) \subseteq N_c$. If $\pi(c_1) = \{c_0, c_1\}$, then there is a coalition of size at least 2 consisting of agents in $C \setminus \{c_0, c_1\}$, and c_1 could perform an MOS deviation to join them. Hence, using Claim 6, it follows that $\pi(c_1) \subseteq C \setminus \{c_0\}$.

Let $\Phi \subseteq C \setminus \{c_0\}$ be a coalition of largest size. Note that $C \setminus \{c_0\}$ cannot contain (at least) 2 singleton coalitions. Then, the singleton with the lower index would join the other singleton. If $|\Phi| = 2$, then $C \setminus \{c_0\}$ consists of two pairs and c_1 has an MOS deviation to join the other pair. Next, assume that $|\Phi| = 3$. If c_1 or c_2 remain as a singleton, they would join Φ. If c_3 or c_4 remain as a singleton, then c_2 performs an MOS deviation to join her. This leaves only the case $|\Phi| = 4$ and we can conclude that $C \setminus \{c_0\} \in \pi$. But then, by $(**)$, z_k has an MOS deviation to join $C \setminus \{c_0\}$ for $k \in [3]$, a contradiction. This concludes Case 1.

Case 2: It holds that $|N_z \cap \Gamma| = 2$.. Then, $|\Gamma \setminus N_z| \leq 2$ which means that $\Gamma \setminus N_z = \{a_i, b_j\}$ and it follows that $\Gamma \cap N_c = \Gamma \cap N_d = \emptyset$. Let $k^* \in [3]$ be the unique index with $z_{i^*} \notin \Gamma$, say without loss of generality $k^* = 1$. Using $(*)$, it must also be the case that $\pi(z_1) \cap N_c = \emptyset$ or $\pi(z_1) \cap N_d = \emptyset$, say without loss of generality $\pi(z_1) \cap N_c = \emptyset$. The identical arguments as in the previous case show that $C \setminus \{c_0\} \in \pi$. But then z_3 could perform an MOS deviation to join $C \setminus \{c_0\}$, a contradiction. This concludes Case 2 and therefore the proof of the claim. $\triangleleft$

Claim 8. *There exists no $\Gamma \in \pi$ with $N_z \subseteq \Gamma$.*

Proof. Assume for contradiction that there exists $\Gamma \in \pi$ with $N_z \subseteq \Gamma$. By Claim 6 and Claim 7, there exists $x \in \{a, b, c, d\}$ with $\Gamma \subseteq N_z \cup N_x$. Without loss of generality, assume that $\Gamma \subseteq N_z \cup N_a$. By Claim 6, $a_0 \notin \Gamma$. We claim that $|\Gamma \cap N_a| \leq 3$. For the contrary, assume that $|\Gamma \cap N_a| = 4$. Then, Claim 6 implies that $\{a_0\} \in \pi$. Also, $v_{a_1}(\pi) = 0$ and $|F_{\text{in}}(\Gamma, a_1)| = |N_z| = |\{a_2, a_3, a_4\}| = |F_{\text{out}}(\Gamma, a_1)|$. Hence, a_1 can perform an MOS deviation to join $\{a_0\}$, a contradiction. Thus, $|\Gamma \cap N_a| \leq 3$, as claimed.

As in the proof of Claim 7, we can show that $B \setminus \{b_0\} \in \pi$. But then z_k has an MOS deviation to join this coalition for every $k \in [3]$, a contradiction. This concludes the proof of this claim. $\lhd$

We are ready to obtain a final contradiction. By Claim 8, there exist $i, j \in [3]$ with $z_i \notin \pi(z_j)$. Without loss of generality, we may assume that $i = 2$ and $j = 1$.

Case 1: It holds that $z_3 \in \pi(z_2)$.. By Claim 6, $v_{z_k}(x) = 1$ for all $k \in [3], x \in (\pi(z_1) \cup \pi(z_2)) \setminus N_z$. Let $m_1 = |\pi(z_2)| - 2 = |\pi(z_2) \setminus N_z|$ and $m_2 = |\pi(z_1)| - 1 = |\pi(z_1) \setminus N_z|$.

If $m_2 \geq m_1$, then z_3 can perform an NS deviation to join $\pi(z_1)$. This is also an MOS deviation unless $\pi(z_2) = \{z_2, z_3\}$. But in this case we find a coalition of the form $N_x \setminus \{x_0\}$ for some $x \in \{a, b, c, d\}$ as in the previous steps. Then, z_2 has an MOS deviation to join this coalition.

On the other hand, if $m_2 < m_1$, then z_1 can perform an MOS deviation to join $\pi(z_2)$. This concludes Case 1. By symmetry, this covers even all cases where at least two agents from N_z are in the same coalition. Hence, it remains one final case.

Case 2: The agents in N_z are in pairwise disjoint coalitions.. Let $p_k = |\pi(z_k)|$ for $k \in [3]$ and $k^* = \arg\max_{k \in [3]} p_i$. Without loss of generality, $k^* = 1$. As in the previous case, it follows from Claim 6 that $v_{z_k}(x) = 1$ for all $k \in [3], x \in \bigcup_{l \in [3]} \pi(z_l) \setminus N_z$. But then z_3 has an MOS deviation to join $\pi(z_1)$. This is the final contradiction and completes the proof. $\square$

Towards the hardness reduction, we start with a useful lemma. It lets us separate the agent set into subsets such that agents from different subsets cannot form joint coalitions within an MOS partition.

Lemma 5. *Consider an FEG (N, v) with an MOS partition π. Let $i, j \in N$ be two agents with $v_i(j) = v_j(i) = -1$ and assume that, for every agent $k \in N \setminus \{i, j\}$, it holds that*

- $v_i(k) = -1$ *or* $v_j(k) = -1$,
- $v_k(i) = -1$ *or* $v_k(j) = -1$,
- $v_k(i) = -1$ *whenever* $v_j(k) = 1$, *and*
- $v_k(j) = -1$ *whenever* $v_i(k) = 1$.

Then, $i \notin \pi(j)$.

Proof. Let an FEG (N, v) be given together with an MOS partition π, and let $i, j \in N$ be two agents satisfying the assumptions of the lemma. Assume for contradiction that $i \in \pi(j)$, and define $C := \pi(j)$. We will use the first assumption of the lemma to show that either i or j can perform an NS deviation to form a singleton coalition, and the other conditions to argue that there is even a valid MOS deviation. First, note that the first assumption implies that, for every agent $k \in N \setminus \{i,j\}$, it holds that $v_i(k) + v_j(k) \leq 0$. Hence,

$$v_i(\pi) + v_j(\pi) = -2 + \sum_{k \in \pi(j) \setminus \{i,j\}} v_i(k) + v_j(k) \leq -2.$$

Therefore, $v_i(\pi) < 0$ or $v_j(\pi) < 0$, and thus either i or j can perform an NS deviation to form a singleton coalition.

In addition, our second assumption implies that, for every agent $k \in N \setminus \{i,j\}$, it holds that $k \in F_{\text{out}}(C,i)$ or $k \in F_{\text{out}}(C,j)$. Hence, a similar averaging argument as the previous consideration shows that $|F_{\text{out}}(C,i)| > |C|/2$ or $|F_{\text{out}}(C,j)| > |C|/2$.

Assume first that $v_i(\pi) < 0$ and $v_j(\pi) < 0$. Then, our second observation implies that one of i and j can perform an MOS deviation to form a singleton coalition, a contradiction. Hence, we may assume without loss of generality that $v_i(\pi) < 0$ and $v_j(\pi) \geq 0$. Then,

$$|F_{\text{in}}(C,i)| - |F_{\text{out}}(C,i)| = |\{l \in C \setminus \{i\} \colon v_l(i) = 1\}| - |\{l \in C \setminus \{i\} \colon v_l(i) = -1\}|$$
$$\leq |\{l \in C \setminus \{i\} \colon v_j(l) = -1\}| - |\{l \in C \setminus \{i\} \colon v_j(l) = 1\}| = -v_j(\pi) \leq 0.$$

In the inequality, we have used the third assumption of the lemma (the forth assumption is necessary for the symmetric case where i and j swap roles). Hence, agent i can perform an MOS deviation to form a singleton coalition. This is a contradiction and we can conclude that $i \notin \pi(j)$. □

We proceed with proving the hardness result.

Theorem 10. *Deciding whether an FEG contains an MOS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let (R,S) be an instance of E3C. We define a reduced FEG (N,v) as follows. By Proposition 6, there exists an FEG without an MOS partition and we may assume that (N',v') is such an FEG with the additional property that there exists an agent $x \in N'$ such that the FEG restricted to $N' \setminus \{x\}$ contains an MOS partition π'. Indeed, an FEG with the additional property can be obtained simply by removing agents until the property is satisfied.

Now, let $N = N_R \cup N_S$ where $N_R = \cup_{r \in R} N^r$ with $N^r = \{y^r \colon y \in N'\}$ for $r \in R$ and $N_S = \cup_{s \in S} N_s$ with $N_s = \{s_0\} \cup \{s_r \colon r \in s\}$ for $s \in S$. Specifically, we denote the agent corresponding to the special agent $x \in N'$ by x^r. Agents of the type s^r will receive a positive utility from forming a coalition with x^r and therefore have the capability of forcing x^r to stay in a coalition of size 2 with them.

We define utilities v as follows:

- For all $s \in S$, $y, z \in N_s$: $v_y(z) = 1$.
- For all $s \in S$, $r \in s$: $v_{s_r}(x^r) = 1$.

- For all $r \in R$ and $y, z \in N'$: $v_{y^r}(z^r) = v'_y(z)$, i.e., the internal valuations for agents in N^r are identical to the valuations in the counterexample (N', v').
- All other valuations are -1.

We claim that (R, S) is a Yes-instance if and only if the reduced FEG contains an MOS partition.

$\implies$. Suppose first that $S' \subseteq S$ partitions R. We define a partition π as follows.

- For $s \in S \setminus S'$: $N_s \in \pi$ and for $s \in S'$: $\{s_0\} \in \pi$.
- For $s \in S'$, $r \in s$: $\{s_r, x^r\} \in \pi$.
- For $r \in R$ and $z \in N' \setminus \{x\}$: $\pi(z^r) = \{y^r \in y \in \pi'(x)\}$.

We claim that the partition π is MOS.

- Let $r \in R$ and consider an agent $y \in N' \setminus \{x\}$. Then, y^r cannot perform an MOS deviation to join $\pi(z^r)$ for any $z \in N' \setminus \{x\}$, because π' restricted to $N' \setminus \{x\}$ is an MOS partition. Moreover, joining $\pi(z)$ for any $z \in N \setminus N^r$ yields utility at most 0 (in fact, the only such coalition that y^r could join to obtain a utility of 0 is $\pi(x^r)$). Hence, if this constituted an MOS deviation, then forming a singleton coalition would also be an MOS deviation, contradicting the fact that π' is an MOS partition.
- Let $r \in R$. Then, x^r is not allowed to leave her coalition by means of an MOS deviation.
- Let $s \in S'$. Then $v_{s_0}(\pi) = 0$ and joining any other coalition yields utility at most 0. In particular, $v_{s_0}(\pi(s_r) \cup \{s_0\}) = 0$ for all $r \in s$. Moreover, for $r \in s$, $v_{s_r}(\pi) = 1$ and joining any other coalition yields utility at most 1. In particular, $v_{s_r}(\pi(s_0) \cup \{s_r\}) = 1$.
- Let $s \in S \setminus S'$. Then, $\pi(s_0)$ is s_0's most preferred coalition and she has no incentive to perform an MOS deviation. Moreover, for $r \in s$, $v_{s_r}(\pi) = 3$ and joining any other coalition yields a utility of at most 0.

Together, we have shown that π is an MOS partition.

$\impliedby$. For the reverse implication, assume that π is an MOS partition for the reduced instance (N, v). We start by determining the coalitions of agents of the type s_0.

Claim 9. *Let $s \in S$. Then, $\pi(s_0) = \{s_0\}$ or $\pi(x) \subseteq N_s$ for all $x \in N_s$.*

Proof. Let $s \in S$ and define $C := \pi(s_0)$. A close inspection of the utilities in the definition of the reduced instance lets us apply Lemma 5 multiple times to conclude that

- for all $s' \in S \setminus \{s\}$, $C \cap N_{s'} = \emptyset$,

- for all $r \in R \setminus s$, $C \cap N^r = \emptyset$, and
- for all $r \in s$, $C \cap N^r \subseteq \{x^r\}$.

Together, $C \subseteq N_s \cup \{x^r : r \in s\}$. Even more, for $r \in s$, if $x^r \in C$, then $v_{x^r}(\pi) < 0$. In addition, $F_{\text{in}}(C, x^r) \subseteq \{s_r\}$ and $s_0 \in F_{\text{out}}(C, x^r)$. Hence, x^r could perform an MOS deviation to form a singleton coalition. We can therefore conclude that $C \subseteq N_s$.

Assume that $C \supsetneq \{s_0\}$. If $|C| = 3$, then there exists a unique $r \in s$ with $s_r \notin C$. Since s_r has only one friend outside C, this would imply that $v_{s_r}(\pi) \leq 1$ whereas $v_{s_r}(C \cup \{s_r\}) = 3$ and $F_{\text{in}}(\pi(s_r), s_r) = \emptyset$. Hence, s_r could perform an MOS deviation to join C, a contradiction. Hence, $|C| = 2$ or $|C| = 4$. As the latter case corresponds to the situation that $C = N_s$, we only need to consider the former case.

Suppose that $s = \{r_1, r_2, r_3\}$ and that $C = \{s_0, s_{r_1}\}$. Then, it holds that $s_{r_3} \in \pi(s_{r_2})$, as otherwise $v_{s_{r_2}}(\pi) \leq 1$ whereas $v_{s_r}(C \cup \{s_{r_2}\}) = 3$ and $F_{\text{in}}(\pi(s_{r_2}), s_{r_2}) = \emptyset$. But then, $\pi(s_{r_2}) = \{s_{r_2}, s_{r_3}\}$. Any other agent would only have enemies in $\pi(s_{r_2})$ and is allowed to leave by a weak majority. This concludes the proof of the claim. ◁

Our next claim investigates elements $s \in S$ for which $\{s_0\} \in \pi$.

Claim 10. *Let $s \in S$ such that $\{s_0\} \in \pi$. Then, for every $r \in s$, it holds that $\pi(s_r) = \{s_r, x^r\}$.*

Proof. Let $s \in S$ with $\{s_0\} \in \pi$ and consider any $r \in s$. Define $C := \pi(s_r)$ and assume for contradiction that there exists $r' \in s$ with $r' \neq r$ and $s_{r'} \in C$. We can combine the following observations:

- Claim 9 shows that $s'_0 \notin C$ for every $s' \in S \setminus \{s\}$.
- Let $\hat{r} \in R$. We can apply Lemma 5 for s_r (or $s_{r'}$) and an agent in $N^{\hat{r}}$ to show that $C \cap N^{\hat{r}} = \emptyset$ if $\hat{r} \neq r$ (or if $\hat{r} = r$).
- Let $s' \in S$ and $\hat{r} \in s'$. We can apply Lemma 5 for s_r (or $s_{r'}$) and $s'_{\hat{r}}$ to show that $s'_{\hat{r}} \notin C$ if $\hat{r} \neq r$ (or $\hat{r} = r$).

Together, the observations show that $C \subseteq N_s$. But then, $v_{s_0}(C \cup \{s_0\}) \geq 2$, and s_0 could perform an MOS deviation to join C. This is a contradiction and we can conclude that $C \cap N_s = \{r_s\}$.

This means in particular, that $F_{\text{in}}(C, s_r) = \emptyset$. Since $v_{s_r}(\{s_0, s_r\}) = 1$, it must hold that $v_{s_r}(\pi) = 1$. Since the unique friend of s_r outside N_s is x^r, we can conclude that $\pi(s_r) = \{s_r, x^r\}$. ◁

We are ready to finish the proof. Therefore, let $S' := \{s \in S : \{s_0\} \in \pi\}$. We show that S' partitions R in two steps. First, the sets in S' are disjoint. Indeed, if $s, s' \in S'$ with $s \neq s'$ and $r \in s \cap s'$, then Claim 10 implies that $\{s_r, x^r\} \in \pi$ and $\{s'_r, x^r\} \in \pi$, contradicting the fact that π is a partition.

It remains to show that all elements of R are covered by a set in S'. Therefore, consider an arbitrary $r \in R$ and let $y \in N'$. By Lemma 5, $\pi(y^r) \cap N^{r'} = \emptyset$ for all

$r' \in R$ with $r' \neq r$. Moreover, Claim 9 and Claim 10 imply that $\pi(y^r) \cap N_s = \emptyset$ for all $s \in S$ with $r \notin s$. Assume for contradiction that there exists no $s \in S'$ with $r \in s$. Then, Claim 9 implies that $\pi(y^r) \cap N_s = \emptyset$ for all $s \in S$ with $r \in s$. Together, $\pi(y^r) \subseteq N^r$. This means that π restricted to the agents in N^r is an MOS partition, contradicting the fact that such a partition does not exist. Hence, r must be covered by some set in S'. □

Now, we provide the full proof for investigating the FEG without MIS partition. First, we prove a useful lemma showing that certain agents in cliques of mutual friendship playing identical roles have to be in joint coalitions in every MIS partition. This will concern the agents in the sets K_i and B_i^j for $i, j \in [5]$ (cf. Figure 8).

Lemma 6. *Consider an FEG (N, v) with an MIS partition π. Let $W \subseteq N$ such that the following conditions hold:*

1. *For all $i, j \in W$, $k \in N$: $v_i(j) = 1$.*
2. *For all $i, j \in W$, $k \in N$: $v_i(k) = v_j(k)$.*
3. *For all $i \in W$, $k \in N$: $v_i(k) = 1$ implies $v_k(i) = 1$.*

Then, there exists a coalition $C \in \pi$ with $W \subseteq C$.

Proof. Let an FEG (N, v) be given together with an MIS partition π, and let $W \subseteq N$ be a subset of agents that fulfills the three conditions of the assertion. Assume for contradiction that there exist $i, j \in W$ with $\pi(i) \neq \pi(j)$. We may assume without loss of generality that $v_i(\pi) \geq v_j(\pi)$. Consider the deviation where agent j joins $\pi(i)$. Then,

$$v_j(\pi(i) \cup \{j\}) \overset{(1),(2)}{=} 1 + v_i(\pi) > v_j(\pi).$$

Hence, this constitutes an NS deviation. Moreover, since π is MIS, it holds that $v_i(\pi) \geq 0$ and therefore, because the game is an FEG,

$$|\{x \in \pi(i) \setminus \{i\} \colon u_i(x) = 1\}| \geq |\{x \in \pi(i) \setminus \{i\} \colon u_i(x) = -1\}|. \qquad (*)$$

It follows that

$$\begin{aligned}
|F_{\text{in}}(\pi(i), j)| &\overset{(1)}{=} |\{x \in \pi(i) \setminus \{i\} \colon u_x(j) = 1\}| + 1 \\
&\overset{(3)}{\geq} |\{x \in \pi(i) \setminus \{i\} \colon u_j(x) = 1\}| + 1 \overset{(2)}{=} |\{x \in \pi(i) \setminus \{i\} \colon u_i(x) = 1\}| + 1 \\
&\overset{(*)}{\geq} |\{x \in \pi(i) \setminus \{i\} \colon u_i(x) = -1\}| + 1 \overset{(2)}{=} |\{x \in \pi(i) \setminus \{i\} \colon u_j(x) = -1\}| + 1 \\
&\overset{(3)}{\geq} |\{x \in \pi(i) \setminus \{i\} \colon u_x(j) = -1\}| + 1 = |F_{\text{out}}(\pi(i), j)| + 1 > |F_{\text{out}}(\pi(i), j)|.
\end{aligned}$$

Hence, this is even an MIS deviation, a contradiction. □

Proposition 8. *There exists an FEG without an MIS partition.*

Proof. We define an FEG for which we prove that it does not contain an MIS partition. As discussed before, the game is rather large (the number of agents is 183), but it has a lot of structure and an illustration was already provided in Figure 8. Formally, the agent set is given by $N = Z \cup \bigcup_{i\in[5]}(A_i \cup B_i \cup K_i)$, where the exact definitions and interpretation of the subsets of agents is as follows:

- The set of agents $Z = \{z_1, z_2, z_3\}$ forms a directed triangle.
- For $i \in [5]$, the sets $A_i = \{a_i^j : j = \{0, 1, \dots, 9\}$ form cliques which are liked by agents in Z, except for the special agent a_i^0. In turn, all of them like the agents in Z.
- For $i \in [5]$, the sets $K_i = \{k_i^j : j \in [11]\}$ form cliques not liked by agents in Z, but a_i^0 likes these agents.
- For $i \in [5]$, define $B_i = \bigcup_{j=1}^5 B_i^j$, where $B_i^j = \{b_i^{j,l} : l \in [3]\}$. The set B_i^j forms a small clique which tries to tempt agent a_i^j to join if B_i^j is a coalition.

The utilities are defined as

- $v_x(y) = 1$ if $(x, y) \in \{(z_1, z_2), (z_2, z_3), (z_3, z_1)\}$,
- $v_{z_i}(a_j^l) = 1$ if $i \in [3]$, $j \in [5]$, and $l \in [9]$,
- $v_{a_i^j}(a_i^l) = 1$ if $i \in [5]$, $j, l \in \{0, 1, \dots, 9\}$,
- $v_{a_i^j}(z_l) = 1$ if $i \in [5]$, $j \in \{0, 1, \dots, 9\}$, and $l \in [3]$,
- $v_{a_i^0}(k_i^j) = v_{k_i^j}(a_i^0) = 1$ if $i \in [5]$, $j \in [11]$,
- $v_{a_i^j}(b_i^{j,l}) = 1$ if $i, j \in [5]$, $l \in [3]$,
- $v_{b_i^{j,l}}(b_i^{j,l'}) = 1$ if $i, j \in [5]$, $l, l' \in [3]$,
- $v_{k_i^j}(k_i^l) = 1$ if $i \in [5]$, $j, l \in [11]$, and
- all other valuations are -1.

Assume for contradiction that π is an MIS partition for this game. The following observation is helpful in various places:

$$\text{Every agent receives non-negative utility in } \pi\text{, i.e., } v_i(\pi) \geq 0 \text{ for all } i \in N. \tag{$*$}$$

The observation is true because every agent of negative utility could perform an MIS deviation to form a singleton coalition. We will now derive a contradiction proving several claims. The first one is a direct application of Lemma 6 for the agents in sets K_i for $i \in [5]$.

Claim 11. *For all $i \in [5]$, there exists $C \in \pi$ with $K_i \subseteq C$.*

The next claim improves upon the previous claim.

Claim 12. *If $i \in [5]$, then $K_i \in \pi$ or $K_i \cup \{a_i^0\} \in \pi$.*

Proof. Let $i \in [5]$ and assume for contradiction that there exists $C \in \pi$ with $K_i \subseteq C$ and $C \setminus (K_i \cup \{a_i^0\}) \neq \emptyset$. By $(*)$, $v_{k_i^1}(\pi) \geq 0$ and therefore $|C \setminus (K_i \cup \{a_i^0\})| \leq |K_i \cup \{a_i^0\}| - 1 = 11$. Let $x \in C \setminus (K_i \cup \{a_i^0\})$. Then, $a_i^0 \in C$, $|C \setminus (K_i \cup \{a_i^0\})| = 11$, and $v_x(y) = 1$ for all $y \in C \setminus (K_i \cup \{a_i^0\})$. Otherwise, x has at most 10 friends leading to $v_x(\pi) \leq 10 - |K_i| < 0$, contradicting $(*)$. Consequently, the agents $C \setminus (K_i \cup \{a_i^0\})$ form a set of 11 mutual friends which all have a_i^0 as a friend. Such a set of agents does not exist, and we derive a contradiction. ◁

The next two claims make similar structural observations for the agent sets B_i^j. First, we can apply Lemma 6 again for a statement analogous to Claim 11.

Claim 13. *For all $i, j \in [5]$, there exists $C \in \pi$ with $B_i^j \subseteq C$.*

We also refine this claim.

Claim 14. *If $i, j \in [5]$, then $B_i^j \in \pi$ or $B_i^j \cup \{a_i^j\} \in \pi$.*

Proof. Let $i, j \in [5]$ and assume for contradiction that there exists $C \in \pi$ with $B_i^j \subseteq C$ and $C \setminus (B_i^j \cup \{a_i^j\}) \neq \emptyset$. If $|C \setminus (B_i^j \cup \{a_i^j\})| < 3 = |B_i^j|$, then $x \in C \setminus (B_i^j \cup \{a_i^j\})$ has a negative utility, contradicting $(*)$. If $|C \setminus (B_i^j \cup \{a_i^j\})| > 3$, then $b_i^{j,1}$ has negative utility, contradicting $(*)$. Hence, $|C \setminus (B_i^j \cup \{a_i^j\})| = 3$. Moreover, then $a_i^j \in C$ as an agent in $C \setminus (B_i^j \cup \{a_i^j\})$ would have negative utility, otherwise. For similar reasons, the agents in $C \setminus (B_i^j \cup \{a_i^j\})$ have to form a clique of friends of a_i^j.

We will exclude all possible agents in $C \setminus (B_i^j \cup \{a_i^j\})$. First note that the structure we obtained so far holds for arbitrary i and j. Hence, if $a_i^{j'} \in C$ for $j' \in [5] \setminus \{j\}$, then the assertion of Claim 14 is already true for i and j' and therefore $B_i^{j'} \in \pi$. But then, $a_i^{j'}$ can perform an MIS deviation to join $B_i^{j'}$, a contradiction. Thus, since the agents in Z are no mutual friends, there exist $l, l' \in \{6, 7, 8, 9\}$ with $a_i^l \in C$ and $a_i^{l'} \notin C$. By $(*)$, $v_{a_i^{l'}}(\pi) \geq 0$. Moreover, since a_i^l and $a_i^{l'}$ have the identical friends in $N \setminus \{a_i^l, a_i^{l'}\}$ and $a_i^{l'}$ is also a friend of a_i^l, it holds that $v_{a_i^l}(\pi(a_i^{l'}) \cup \{a_i^l\}) \geq 1$. Since $v_{a_i^l}(\pi) = 0$, this is an NS deviation. Also, since all friends of $a_i^{l'}$ and $a_i^{l'}$ herself favor a_i^l to join their coalition, this is even an MIS deviation. Hence, we obtain a contradiction. ◁

The next claim establishes a relationship between agents in Z and A_i.

Claim 15. *For $i \in [5]$, if $Z \cap \pi(a_i^j) = \emptyset$ for all $j \in [9]$, then $A_i \setminus \{a_i^0\} \in \pi$.*

Proof. Let $i \in [5]$ such that $Z \cap \pi(a_i^j) = \emptyset$ for all $j \in [9]$. First, we show that then $\pi(a_i^j) \subseteq A_i$ for $j = 6,7,8,9$. Let therefore $j \in \{6,7,8,9\}$ and assume for contradiction that $\pi(a_i^j) \setminus A_i \neq \emptyset$. By Claim 12, Claim 14, and the initial assumptions of this claim, $\pi(a_i^j) \subseteq \bigcup_{l\in[5]} A_l$. Consider $x \in \pi(a_i^j) \setminus A_i$. If $|\pi(a_i^j) \setminus A_i| \leq |\pi(a_i^j) \cap A_i|$, then $v_x(\pi) < 0$, contradicting $(*)$. On the other hand, if $|\pi(a_i^j) \setminus A_i| \geq |\pi(a_i^j) \cap A_i|$, then $v_{a_i^j}(\pi) < 0$, also contradicting $(*)$. We derived a contradiction in both cases and can therefore conclude that $\pi(a_i^j) \subseteq A_i$.

As in previous steps, we can use the symmetry of the agents in $\{a_i^j : j = 6,7,8,9\}$ to show that there exists a coalition $C \in \pi$ with $\{a_i^j : j = 6,7,8,9\} \subseteq C \subseteq A_i$. Indeed, otherwise, one of these agents could join the coalition of another such agent of at least as high utility by an MIS deviation. Hence, $B_i^j \cup \{a_i^j\} \notin \pi$ for $j \in [5]$ as a_i^j would perform an MIS deviation to join C, otherwise. But then, similarly as above, $\pi(a_i^j) \subseteq A_i$ for $j \in [5]$, and therefore even $A_i \setminus \{a_i^0\} \subseteq C$. Finally, if $a_0^i \in C$, then $v_{a_i^0} = 9$. However, by Claim 12, $K_i \in \pi$ and therefore a_i^0 could perform an MIS deviation to join K_i. Hence, $C = A_i \setminus \{a_i^0\}$. ◁

We have now collected enough structural results to consider the agents in Z. The next two claims will yield the final contradiction.

Claim 16. *There exists no $C \in \pi$ with $Z \subseteq C$.*

Proof. Assume for contradiction that there exists $C \in \pi$ with $Z \subseteq C$. By Claim 12 and Claim 14, $C \subseteq Z \cup \bigcup_{i\in[5]} A_i$. Define $I = \{i \in [5] : A_i \cap C \neq \emptyset\}$ and let

$$i^* \in \arg\min_{i\in I}\{|A_i \cap C|\}. \tag{$**$}$$

Let $x \in A_{i^*} \cap C$.

<u>Case 1</u>: $|I| = 5$.. In this case, we obtain a contradiction to $(*)$ because

$$\begin{aligned} v_x(\pi) &= 3 + (|A_{i^*} \cap C| - 1) - \sum_{i\in I\setminus\{i^*\}} |A_i \cap C| \\ &\overset{(**)}{\leq} 2 - (|I| - 2)|A_{i^*} \cap C| \leq -1 < 0. \end{aligned}$$

<u>Case 2</u>: $|I| = 4$.. As in the previous case, $0 \overset{(*)}{\leq} v_x(\pi) \leq 2 + |A_{i^*} \cap C| - \sum_{i\in I\setminus\{i^*\}} |A_i \cap C|$. Therefore,

$$3|A_{i^*} \cap C| \leq \sum_{i\in I\setminus\{i^*\}} |A_i \cap C| \leq 2 + |A_{i^*} \cap C|.$$

Consequently, $|A_{i^*} \cap C| = 1$ and $|A_i \cap C| = 1$ for $i \in I \setminus \{i^*\}$. Let $l \in [3]$. Then, $v_{z_l}(\pi) \leq 4$. By Claim 15, it holds that $A_{i'} \setminus \{a_{i'}^0\} \in \pi$, where $i' \in [5] \setminus I$. Hence, z_l has an MIS deviation, a contradiction.

<u>Case 3</u>: $|I| = 3.$. As in Case 2,

$$2|A_{i^*} \cap C| \leq \sum_{i \in I \setminus \{i^*\}} |A_i \cap C| \leq 2 + |A_{i^*} \cap C|.$$

Hence, $|A_{i^*} \cap C| \leq 2$ and thus $\sum_{i \in I \setminus \{i^*\}} |A_i \cap C| \leq 4$. Therefore, $v_{z_l}(\pi) \leq 6$ if $l \in [3]$, and an analogous MIS deviation is possible as in the previous case.

<u>Case 4</u>: $|I| = 2.$. Let $i' \in I \setminus \{i^*\}$ be the unique second index in I. We claim that $a_i^j \notin C$ for $i \in I$ and $j \in [5]$. Let $j \in [5]$. First, if $a_{i^*}^j \in C$, then $v_{a_{i^*}^j}(\pi) \leq 3 + (|A_{i^*} \cap C| - 1) - |A_{i'} \cap C| \leq 2$. Moreover, by Claim 14, $B_{i^*}^j \in \pi$ and $a_{i^*}^j$ could perform an MIS deviation to join $B_{i^*}^j$.

Second, assume that $a_{i'}^j \in C$. Then, again by Claim 14, $B_{i'}^j \in \pi$ and since π is MIS, $u_{a_{i'}^j}(\pi) \geq 3$. Let $j' \in [9] \setminus \{j\}$ and assume for contradiction that $a_{i'}^{j'} \notin C$. Since $a_{i'}^{j'}$ has at least as many friends in C as $a_{i'}^j$ (recall that $B_{i'}^j \in \pi$), $v_{a_{i'}^{j'}}(\pi) \geq v_{a_{i'}^j}(\pi) + 1 \geq 4$. Using Claim 14, this means in particular that $B_{i'}^{j'} \cap \pi(a_{i'}^{j'}) = \emptyset$ if $j' \in [5]$. Therefore, $v_{a_{i'}^{j'}}(\pi(a_{i'}^j) \cup \{a_{i'}^{j'}\}) \geq v_{a_{i'}^j}(\pi) + 1$ and $v_{a_{i'}^j}(\pi(a_{i'}^{j'}) \cup \{a_{i'}^j\}) \geq v_{a_{i'}^{j'}}(\pi) + 1$. Hence, either $a_{i'}^{j'}$ has an MIS deviation to join $\pi(a_{i'}^j)$ or $a_{i'}^j$ has an MIS deviation to join $\pi(a_{i'}^{j'})$, a contradiction. Consequently, $a_{i'}(j') \in C$ and therefore $A_{i'} \setminus \{a_{i'}^0\} \subseteq C$.

Recall that we already know that $|A_{i^*} \cap C| \leq 5$ because $a_{i^*}^l \notin C$ for $l \in [5]$. We obtain a contradiction to $(*)$ because

$$v_x(\pi) \leq 3 + (\underbrace{|A_{i^*} \cap C|}_{\leq 5} - 1) - \underbrace{|A_{i'} \cap C|}_{\geq 9} \leq -2 < 0.$$

Hence, we can conclude that $a_{i'}^j \notin C$ for $j \in [5]$. But then, for $l \in [3]$, $v_{z_l} \leq |(A_{i^*} \setminus \{a_{i^*}^0\}) \cap C| + |(A_{i'} \setminus \{a_{i'}^0\}) \cap C| \leq 8$. Hence, z_l can perform an MIS deviation to join $A_i \setminus \{a_i^0\}$ for $i \in [5] \setminus I$, as in the previous two cases.

<u>Case 5</u>: $|I| = 1.$. If $C \neq Z \cup (A_{i^*} \setminus \{a_{i^*}^0\})$, then, for $l \in [3]$, $v_{z_l}(\pi) \leq 8$, and an analogous MIS deviation as in the previous cases is possible. Hence, $C = Z \cup (A_{i^*} \setminus \{a_{i^*}^0\})$. But then $v_{a_{i^*}^0}(\pi) \leq 11$, whereas $v_{a_{i^*}^0}(C \cup \{a_{i^*}^0\}) \geq 12$. Hence, $a_{i^*}^0$ has an MIS deviation to join C (which is favored by all agents in $A_{i^*} \setminus \{a_{i^*}^0\}$). This is a contradiction, and concludes the proof of the claim. $\lhd$

For a final contradiction, it remains to lead the case to a contradiction that the agents in Z are part of different coalitions.

Claim 17. *There exists $C \in \pi$ with $Z \subseteq C$.*

Proof. Assume for contradiction that there exists $C \in \pi$ with $Z \cap C \neq \emptyset$ and $Z \not\subseteq C$.

Assume first that $|Z \cap C| = 2$ and suppose without loss of generality that $z_1, z_2 \in C$. Note that $v_{z_3}(C \cup \{z_3\}) = v_{z_2}(\pi) + 1$. Hence, if $v_{z_3}(\pi) \leq v_{z_2}(\pi)$, then z_3 can perform an NS deviation to join C. This is even an MIS deviation as $v_{z_2}(\pi) \geq 0$ and z_2 favors her to join. On the other hand, $v_{z_2}(\pi(z_3) \cup \{z_2\}) = v_{z_3}(\pi) + 1$. Hence, if $v_{z_2}(\pi) < v_{z_3}(\pi)$, then z_2 has an NS deviation to join $\pi(z_3)$. Note that z_3 is opposed to that. However, as $v_{z_3}(\pi) > v_{z_2}(\pi) \geq 0$, and every friend of z_3 in $\pi(z_3)$ favors to let z_2 join, it holds that

$$\begin{aligned}|F_{\text{in}}(\pi(z_3), z_2)| &= |\{y \in \pi(z_3) \colon u_{z_3}(y) = 1\}| \\ &\geq |\{y \in \pi(z_3) \colon u_{z_3}(y) = -1\}| + 1 \\ &\geq |F_{\text{out}}(\pi(z_3), z_2)|.\end{aligned}$$

Hence, this is even an MIS deviation.

Finally, assume that $\pi(z_l) \cap Z = \{z_l\}$ for all $l \in [3]$. Let $l \in [3]$ and $i \in [5]$. Then, $a_i^0 \notin \pi(z_l)$. Indeed, if $a_i^0 \in \pi(z_l)$, then u_i^0 can have at most 10 friends in her coalition. By Claim 12, $K_i \in \pi$ and a_i^0 would perform an MIS deviation to join this coalition. By this observation and using Claim 12 and Claim 14, z_l forms a coalition with friends only (and these do additionally also have all agents in Z as a friend).

Let $l^* \in \arg\min_{l \in [3]}\{v_{z_l}(\pi)\}$. Without loss of generality, we may assume that $l^* = 1$. Then, z_1 has an NS deviation to join $\pi(z_2)$. This is also an MIS deviation unless $\pi(z_2) = \{z_2\}$. Then, z_2 has an NS deviation to join $\pi(z_3)$, which in turn is an MIS deviation unless $\pi(z_3) = \{z_3\}$. By the minimality assumption on l^*, it must then also hold that $\pi(z_1) = \{z_1\}$. But then, using Claim 15, $A_1 \setminus \{a_1^0\} \in \pi$ and z_1 could perform an MIS deviation to join this coalition. This contradiction concludes the proof of the claim. ◁

As the combination of Claim 16 and Claim 17 directly leads to a contradiction, we have shown that the constructed FEG has no MIS partition. □

Towards turning this counterexample into an intractability result for FEGs, we prove another useful lemma, which excludes that enemies can be in a joint coalition of an MIS partition if they do not have a common friend in their coalition.

Lemma 7. *Consider an FEG (N, v) with an MIS partition π. Let $i, j \in N$ be two agents with $v_i(j) = v_j(i) = -1$ such that, for every agent $k \in \pi(i) \setminus \{i, j\}$, it holds that $v_i(k) = -1$ or $v_j(k) = -1$. Then, $i \notin \pi(j)$.*

Proof. Let an FEG (N, v) be given together with an MIS partition π, let $i, j \in N$ be two agents satisfying the assumptions of the lemma. Assume for contradiction that $i \in \pi(j)$. Our assumptions imply in particular that, for every agent $k \in N \setminus \{i, j\}$, it holds that $v_i(k) + v_j(k) \leq 0$. Hence,

$$v_i(\pi) + v_j(\pi) = -2 + \sum_{k \in \pi(j) \setminus \{i,j\}} v_i(k) + v_j(k) \leq -2.$$

Therefore $v_i(\pi) < 0$ or $v_j(\pi) < 0$, a contradiction. □

Theorem 11. *Deciding whether an FEG contains an MIS partition is NP-complete.*

Proof. We provide a reduction from E3C. Let (R, S) be an instance of E3C. We define an FEG (N, v) as follows. Let $N = N_R \cup N_S$ where $N_R = \cup_{r \in R} N^r$ and $N_S = \cup_{s \in S} N_s$ with $N_s = V_s \cup \bigcup_{r \in s} V_s^r$ for $s \in S$. There, we define, for $s \in S$, $V_s = \{c_{s,i} \colon i \in [10]\}$, and for $s \in S$ and $r \in s$, $V_s^r = \{c_{s,i}^r \colon i \in [10]\}$. To define the sets N^r, assume that (N', v') is the FEG constructed in the proof of Proposition 8. Then, for $r \in R$, we define $N^r = \{x^r \colon x \in N'\}$. Specifically, we denote the agent corresponding to z_1 by z_1^r. Agents of this type will be linked to agents in V_s^r by means of a positive utility correspondence. We define utilities v as follows:

- For all $s \in S$, $x, y \in N_s$: $v_x(y) = 1$.
- For all $s \in S$, $r \in s$, and $x \in V_s^r$: $v_x(z_1^r) = v_{z_1^r}(x) = 1$.
- For all $r \in R$ and $x, y \in N'$: $v_{x^r}(y^r) = v'_x(y)$ i.e., the internal valuations for agents in N^r are identical to the valuations in the counterexample defined in the proof of Proposition 8.
- All other valuations are -1.

We claim that (R, S) is a Yes-instance if and only if the reduced FEG contains an MIS partition.

$\Longrightarrow$. Suppose first that $S' \subseteq S$ partitions R. We define a partition π based on a partition π' of the agent set $N' \setminus \{z_1\}$ in the game (N', v') from the proof of Proposition 8. The partition π' is given as follows.

- We have $\{z_2, z_3\} \cup A_1 \in \pi'$ and $K_1 \in \pi'$.
- For $i, j \in [5]$, $B_i^j \in \pi'$.
- For $i \in \{2, 3, 4, 5\}$, $A_i \setminus \{a_i^0\} \in \pi'$ and $K_i \cup \{a_i^0\} \in \pi'$.

Based on this partition, we can define the partition π as follows.

- For $s \in S \setminus S'$: $N_s \in \pi$ and for $s \in S'$: $V_s \in \pi$.
- For $s \in S'$, $r \in s$: $V_s^r \cup \{z_1^r\} \in \pi$.
- For $r \in R$ and $x \in N' \setminus \{z_1\}$: $\pi(x^r) = \{y^r : y \in \pi'(x)\}$.

Showing that π is MIS follows from a lengthy, but straightforward case analysis.

- For every $r \in R$ and $x \in N' \setminus \{z_1\}$, agent x^r has utility $v_{x^r}(\pi) > 0$, and therefore x^r cannot join a coalition containing an agent outside N^r as this would give her negative utility. Moreover, also deviations within N^r cannot improve her utility:
  - For $i, j \in [5]$, and $l \in [3]$, if $x = b_i^{j,l}$, then $v_{x^r}(\pi) = 2$, but x^r can have at most one friend in any other coalition.

- For $i \in [5]$ and $j \in [11]$, if $x = k_i^j$, then $v_{x^r}(\pi) \geq 10$, but x^r can have at most one friend in any other coalition.
- If $x = a_1^0$, then $v_{x^r}(\pi) = 11$, and the only possible deviation that gives x^r positive utility, i.e., joining K_1, would not increase her utility.
- For $i \in \{2, 3, 4, 5\}$, if $x = a_i^0$, then $v_{x^r}(\pi) = 11$, and the only possible deviation that gives x^r positive utility, i.e., joining $A_i \setminus \{a_i^0\}$ would decrease her utility.
- If $x = z_2$ or $x = z^3$, then $v_{x^r}(\pi) \geq 9$, and the only possible deviations, i.e., joining a coalition $A_i \setminus \{a_i^0\}$ for $i \in \{2, 3, 4, 5\}$ would not increase her utility.

- For $r \in R$, $v_{z_1^r}(\pi) = 10$, and joining any other coalition does not increase her utility.
- For $s \in S \setminus S'$ and $x \in N_s$, $v_x(\pi) = 39$, and joining any other coalition does not give agent x positive utility.
- For $s \in S'$ and $x \in V_s$, $v_x(\pi) = 9$, and joining any other coalition does not give her a better utility. In particular, joining $V_s^r \cup \{z_1^r\}$ for $r \in s$ would also give her a utility of 9.
- For $s \in S'$, $r \in s$, and $x \in V_s^r$, $v_x(\pi) = 10$, and no other coalition gives her a better utility. In particular, joining V_s would also give her a utility of 10.

Together, we have shown that π is an MIS partition (we have even shown that it is an NS partition).

$\Longleftarrow$. Conversely, assume that the reduced FEG contains an MIS partition π.

Note that the assumptions of Lemma 7 are in particular satisfied for two agents $i, j \in N$ with $v_i(j) = v_j(i) = -1$ such that, for *every* agent $k \in N \setminus \{i, j\}$, it holds that $v_i(k) = -1$ or $v_j(k) = -1$. Therefore, we can apply Lemma 7 multiple times to obtain the following facts:

1. For $r, r' \in R$ with $r \neq r'$, $x \in N^r$, and $y \in N^{r'}$, it holds that $y \notin \pi(x)$.
2. For every $s, s' \in S$, $s \neq s'$, $x \in V_s$, and $y \in N_{s'}$, it holds that $y \notin \pi(x)$.
3. For every $s \in S$, $r \in R \setminus s$, $x \in N_s$, and $y \in N^r$, it holds that $y \notin \pi(x)$.
4. For every $s \in S$, $r \in s$, and $x \in V_s$, it holds that $\pi(x) \cap N^r \subseteq \{z_1^r\}$.

Next, we can apply Lemma 6 to obtain the next two facts.

5. For every $s \in S$, there exists a coalition $C \in \pi$ with $V_s \subseteq C$.
6. For every $s \in S$, $r \in S$, there exists a coalition $C \in \pi$ with $V_s^r \subseteq C$.

Moreover, combining Lemma 7 with Fact 6 allows us to further refine Fact 4 yielding the fact

7. For every $s \in S$, $r \in s$, and $x \in V_s$, it holds that $V_s^r \subseteq \pi(x)$ whenever $z_1^r \in \pi(x)$.

We are ready to restrict the coalitions of agents in sets V_s to two possibilities.

Claim 18. *For all $s \in S$, it holds that $V_s \in \pi$ or $N_s \in \pi$.*

Proof. Let $s \in S$ and $x \in V_s$, and define $C := \pi(x)$. By Fact 5, $V_s \subseteq C$. Furthermore, by Fact 2, Fact 3, and Fact 4, it holds that $C \subseteq N_s \cup \{z_1^r : r \in s\}$.

Suppose that $V_s \subsetneq C$. We have to show that $C = N_s$. By Fact 7, there exists $r \in s$ with $V_s^r \subseteq C$. Assume for contradiction that $z_1^r \in C$. Since all agents in C except the agents in N_s^r are enemies of z_1^r, it holds that $v_{z_1^r}(\pi) < 0$ if $C \supsetneq V_s \cup V_s^r \cup \{z_1^r\}$. This would contradict that π is an MIS partition and therefore $C = V_s \cup V_s^r \cup \{z_1^r\}$. In particular, every agent $y \in N_s \setminus C$ has to satisfy $v_y(\pi) \geq 19$. Otherwise, this agent could perform an MIS deviation to join C. Hence, there exists a coalition $D \in \pi$ with $N_s \setminus C \subseteq D$. Assume that $s = \{r, r', r''\}$. Let $y' \in V_s^{r'}$ and $y'' \in V_s^{r''}$. If there exists an agent $q \in N \setminus (V_s^{r'} \cup V_s^{r''})$, then either $v_{y'}(q) = -1$ or $v_{y''}(q) = -1$. Assume without loss of generality that the former case holds. Then, $z_1^{r'} \in D$. Otherwise, $v_{y'}(\pi) \leq 18$ and y' would deviate to join C. But then also $z_1^{r''} \in D$ (due to the utility of y''), and it must hold that $D = V_s^{r'} \cup V_s^{r''} \cup \{z_1^{r'}, z_1^{r''}\}$. But then, $v_{z_1^{r'}}(\pi) = -1$, a contradiction. Hence, $D = V_s^{r'} \cup V_s^{r''}$. but then, any agent in V_s has an MIS deviation to join D, a contradiction. We can conclude that $z_1^r \notin C$.

Since the previous argument is valid for every $r \in s$ with $V_s^r \subseteq C$, we can conclude that $C \subseteq N_s$. Assume for contradiction that there exists an agent $y \in N_s \setminus C$, say without loss of generality that $y \in V_s^{r'}$. Note that $v_y(C \cup \{y\}) \geq 20$, and therefore, it must hold that $v_y(\pi) \geq 20$. Hence, $V_s^{r'} \cup V_s^{r''} \cup \{z_1^{r'}\} \subseteq \pi(y)$. Therefore, even $z_1^{r''} \in \pi(y)$ because otherwise, an agent in $V_s^{r''}$ would perform an MIS deviation to join C. But then, as in the previous argument, $z_1^{r'}$ has a negative utility, a contradiction. Hence, $C = N_s$. This concludes the proof of the claim. ◁

Our next goal is to pinpoint the coalitions of agents in sets of the type V_s^r.

Claim 19. *For all $s \in S$ and $r \in s$, it holds that $V_s^r \cup \{z_1^r\} \in \pi$ or $N_s \in \pi$.*

Proof. For $s \in S$ and $r \in s$ consider an agent $x \in V_s^r$ and define $C := \pi(x)$. Assume that $C \neq N_s$. We have to show that $C = V_s^r \cup \{z_1^r\}$. By Claim 18, we know then that $V_s \in \pi$. By Fact 3, we know that $C \subseteq N_S \cup \bigcup_{t \in s} N^t$. Assume that $s = \{r, r', r''\}$.

Assume for contradiction that there exists an agent $y \in (V_s^{r'} \cup V_s^{r''}) \cap C$. Then, $C \cap N^t \subseteq \{z_1^t\}$ for $t \in s$. Indeed, if there is $t \in s$ and an agent $q \in (N^t \setminus \{z_1^t\}) \cap C$, then we derive a contradiction by applying Lemma 7 for q and one of x and y. A similar argument shows that $N_S \cap C \subseteq N_s$. Hence, $C \subseteq N_s \cup \bigcup_{t \in s} \{z_1^t\}$.

By Fact 6 and our assumptions, we know that in addition $V_s^r \cup V_s^t \subseteq C$ for $t \in s$ with $y \in V_s^t$. Hence, $v_p(C \cup \{p\}) \geq 17 > 9 = v_p(\pi)$ for every $p \in V_s$. Hence, such an agent p could perform an MIS deviation, a contradiction. We can therefore conclude that $C \cap N_s = V_r^s$. Since $V_s \in \pi$, it must hold that $v_x(\pi) \geq 10$.

Since we already know that $C \subseteq N_s \cup (N_S \setminus N_s) \cup \bigcup_{t \in s} N^t$, this is only possible if $C = V_s^r \cup \{z_1^r\}$. $\lhd$

We are ready to prove that (R, S) is a Yes-instance. Define $S' = \{s \in S \colon N_s \notin \pi\}$. First, note that the sets in S' are disjoint. Indeed, let $s \in S'$ and consider $r \in s$. By Claim 19, $V_s^r \cup \{z_1^r\} \in \pi$. Hence, for every $s' \in S \setminus \{s\}$ with $r \in s'$, it cannot be the case that $V_{s'}^r \cup \{z_1^r\} \in \pi$. Hence, another application of Claim 19 yields $N_{s'} \in \pi$, and therefore $s' \notin S'$.

It remains to show that S' covers all elements in R. Therefore, let $r \in R$. By Fact 1, Claim 18, and Claim 19, it holds that $\pi(x) \subseteq N^r$ for all $x \in N^r \setminus \{z_1^r\}$ and $\pi(z_1^r) \subseteq N^r$ or $\pi(z_1^r) = V_s^r \cup \{z_1^r\}$ for some $s \in S$. In the former case, $\pi(x) \subseteq N^r$ for all $x \in N^r$, which contradicts the fact that π is an MIS partition because, according to the proof of Proposition 8, the game restricted to N^r contains no MIS partition. Hence, the latter case must be true, i.e., $\pi(z_1^r) = V_s^r \cup \{z_1^r\}$ for some $s \in S$. Then, $s \in S'$, and therefore r is covered by an element in S'. $\square$

Appendix E. Missing Proof in Section 5.4

Theorem 12. *Deciding whether an ASHG contains an SMS (or JMS) partition is* NP*-complete.*

Proof. We provide a polynomial-time reduction from E3C that simultaneously works for JMS and SMS. Let (R, S) be an instance of E3C. We produce an ASHG (N, v) such that for all $\alpha \in \{\text{JMS, SMS}\}$, (R, S) has an exact cover if and only if (N, v) has a partition that is α. Define the agent set $N = \bigcup_{s \in S} A^s \cup \bigcup_{r \in R} \bigcup_{i=1}^{n_r} B_i^r$, where $A^s = \{a_r^s \colon r \in s\}$ for $s \in S$ and $B_i^r = \left\{b_{i,j}^r \colon j \in [5]\right\}$ for $r \in R, i \in [n_r - 1]$.

Also, define utilities v as follows:

- For each $s \in S, a \neq a' \in A^s : v_a(a') = 2$.
- For each $r \in R, s \in S_r, i \in [n_r - 1] : v_{a_r^s}(b_{i,1}^r) = 1, v_{b_{i,1}^r}(a_r^s) = 0$.
- Each B_i^r has internal utilities as in the example constructed in Proposition 9, i.e., if v' are the utilities in the example, then $v_{b_{i,j}^r}(b_{i,k}^r) = v'_j(k)$.
- All other valuations are $-M$, where $M = |S| + 5$ (can be thought of as $-\infty$).

The reduction is visualized in Figure E.12. Note that the it can be performed in polynomial time, as there are at most $3|S| + 5|R||S|$ agents. We proceed with the proof of the correctness of the reduction and show that if (R, S) has an exact cover, then (N, v) also has a JMS and SMS partition, and conversely if (N, v) has a partition π that is either JMS or SMS, then there is an exact cover in the instance (R, S).

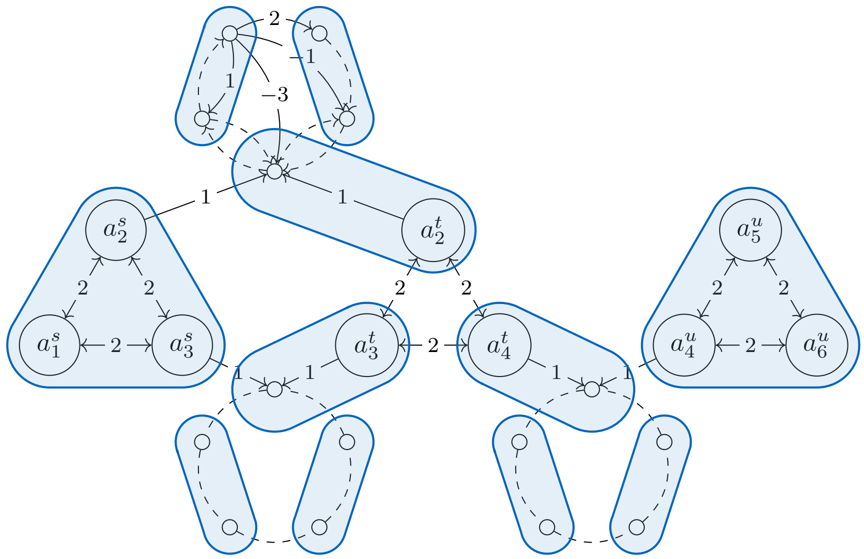


Figure E.12: Schematic of the reduction from the proof of Theorem 12 for the Yes-instance of E3C $(\{1,\dots,6\},\{s,t,u\})$ with $s=\{1,2,3\}, t=\{2,3,4\}$ and $u=\{4,5,6\}$. Some edges have been omitted for clarity. The indicated partition is both SMS and JMS.

$\Longrightarrow$:. Suppose (R,S) has an exact cover $S' \subseteq S$. We construct a stable partition π.

- We have coalitions corresponding to the cover, i.e., for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \dots, s_{n_r-1}$ and define for each $i \in [n_r - 1]$ the coalitions $\left\{a_r^{s_i}, b_{i,1}^r\right\}, \left\{b_{i,2}^r, b_{i,3}^r\right\}, \left\{b_{i,4}^r, b_{i,5}^r\right\}$.

We claim that this partition is JMS and SMS. To see this, note that the only agents that have incentive to deviate are agents of type $b_{i,1}^r$ who would prefer to join $\left\{b_{i,2}^r, b_{i,3}^r\right\}$. Fix any such agent $b_{i,1}^r$. The agent a_r^s she is paired with would vote against her leaving, so the partition is MOS and thus SMS. To see that it is also JMS, note that even though $b_{i,2}^r$ would vote in favor of the deviation, $b_{i,3}^r$ is against it, which together with the against-vote of a_r^s ensures that there is a strict joint majority against the deviation.

$\Longleftarrow$:. Suppose there is a partition π that is JMS or SMS. We show that then there must be an exact cover $S' \subseteq S$ of R. We begin with some observations:

1. Agents $b_{i,j}^r$ with $j \in \{2,\dots,5\}$ must have $\pi(b_{i,j}^r) \subseteq B_i^r$. For contradiction, suppose this is not so. Consider first the case that there is exactly one outside agent $a \in \pi(b_{i,j}^r) \setminus B_i^r$. Then, as $v_a(b_{i,j}^r) = -M$, a has incentive to form a singleton coalition, and this is a valid SMS deviation (and therefore

JMS deviation). The other case is that there are at least two agents $a \neq a' \in \pi(b^r_{i,j}) \setminus B^r_i$. Then, as $v_{b^r_{i,j}}(a) = -M$ and $\left|F_{\text{out}}(\pi(b^r_{i,j}), b^r_{i,j})\right| \geq |\{a, a'\}| = 2 = \left|\left\{b^r_{i,j+_{[5]}1}, b^r_{i,j+_{[5]}4}\right\}\right| \geq \left|F_{\text{in}}(\pi(b^r_{i,j}), b^r_{i,j})\right|$, $b^r_{i,j}$ can form a singleton coalition.

2. Agents a^s_r and $a^{s'}_{r'}$ with $s \neq s'$ have $\pi(a^s_r) \neq \pi(a^{s'}_{r'})$. For contradiction, suppose the contrary, i.e., suppose that there are a^s_r and $a^{s'}_{r'}$ with $s \neq s'$, but $\pi(a^s_r) = \pi(a^{s'}_{r'}) =: C$. As $v_{a^s_r}(a^{s'}_{r'}) = v_{a^{s'}_{r'}}(a^s_r) = -M$, both would rather be in a singleton coalition. Further, we can assume without loss of generality that $|A^s \cap C| \leq \left|A^{s'} \cap C\right|$ (otherwise, we can just swap them). Then, as $|F_{\text{out}}(C, a^s_r)| \geq \left|A^{s'} \cap C\right| \geq |A^s \cap C| > |F_{\text{in}}(C, a^s_r)|$, a^s_r can deviate to form a singleton coalition.

3. Agents $b^r_{i,1}$ must be in a pair with exactly one agent a^s_r. Fix such an agent $b^r_{i,1}$. First, due to observation 1, she cannot be alone, and no other agents from B^r_i can be in her coalition, as the example constructed in Proposition 9 has no SMS partition. Consequently, she must form a coalition with at least one agent outside of B^r_i, and no agents from B^r_i. Next, due to observation 2, she can be together with at most one agent of type a^s_r. If there was another member from A^s (other than a^s_r), $b^r_{i,1}$ could deviate to a singleton coalition.

We now know that for each $r \in R$, exactly $n_r - 1$ of the agents a^s_r must be in pairs with agents $b^r_{i,1}$. This leaves exactly one agent a^s_r not in a pair. We claim that for these agents we have $\pi(a^s_r) = A^s$, yielding a cover $S' = \{s \in S \colon A^s \in \pi\}$. Suppose that a^s_r is such an agent not in a pair. Then, $\pi(a^s_r) \subseteq A^s$. If the other two agents from A^s form a pair, then a^s_r has an incentive to join them. Otherwise, the other two agents would have an incentive to join a^s_r. In any case, the only stable situation is $\pi(a^s_r) = A^s$. □